Historic, archived document

Do not assume content reflects current scientific knowledge, policies, or practices.

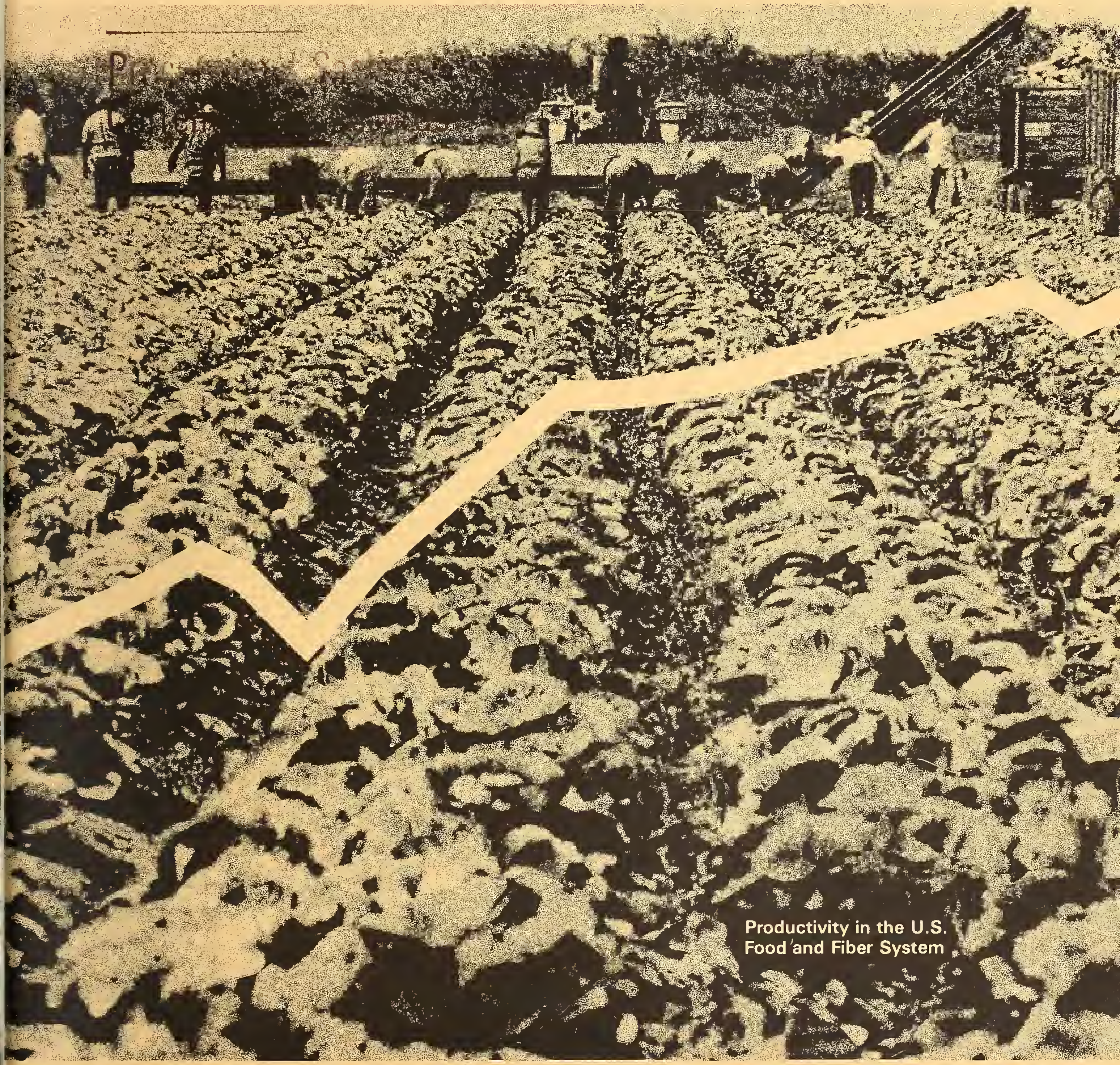
Ec7Agr

AGRICULTURAL ECONOMICS RESEARCH

A JOURNAL OF THE
U.S. DEPARTMENT OF AGRICULTURE
ECONOMIC RESEARCH SERVICE

OCTOBER 1976 VOL. 28 NO. 4

U. S. D. A.
National Agricultural Library
Received



Productivity in the U.S.
Food and Fiber System

AGRICULTURAL ECONOMICS RESEARCH

OCTOBER 1976 VOL. 28 NO. 4

A Journal of Economic and Statistical Research in the United States Department of Agriculture and Cooperating Agencies

Editors
Judith A. Armstrong
Clark Edwards

Graphic Designers
Jan Proctor
James Vechery

Staff Assistant
Ann Bassett

Editorial Board
Charles Caudell
Robert Coltrane
William Kost
Jitendar Mann
Robert McKusick
J.B. Penn
Roberta Van Haeften

Contents *Page*

In This Issue 146

Measuring Labor Productivity in Production of Food for Personal Consumption 123
Eric C. Howe, Gerald E. Schluter, and Charles R. Handy

Productivity Changes in the Food and Fiber System, 1958-74 130
Donald D. Durost and James E. Kirkley

Least Squares Robustness When the Error Term Is Misspecified in Cobb-Douglas Type Functions 136
Terry N. Barr and James F. Horrell

Research Review 146
- The Fear of Knowledge, Charles E. Bishop
- Policies for Equitable Growth in Developing Countries, Irma Adelman
- *Regional Policy: Readings in Theory and Application*, reviewed by Clark Edwards
- A Note on the Recent Decline in the Average Age of Farmers, Calvin Beale
- Technology Assessment Workshop, W. B. Back

MEASURING LABOR PRODUCTIVITY IN PRODUCTION OF FOOD FOR PERSONAL CONSUMPTION

By Eric C. Howe, Gerald E. Schluter, and Charles R. Handy*

Labor productivity is measured for each of 4 selected years for each of four components of personal consumption expenditures for food: food purchased for off-premises consumption, purchased meals and beverages, food furnished Government and commercial employees, and food produced and consumed on farms. The labor productivity estimates include both direct and supporting labor needs because, through input-output analysis, it was possible to identify supporting labor inputs regardless of their sectoral location in the economy.

Keywords: Food and fiber sector analysis, input/output analysis, labor, personal consumption, productivity.

INTRODUCTION

The principal purpose of this article is to report a new measure of labor productivity for U.S. production of food for personal consumption. We measure productivity for the aggregate and for four components.

The orientation of the study differs fundamentally from the norm. Most productivity studies proceed entirely along production-oriented lines. The analysts aggregate activities of specific producers because they produce similar goods and services. Such studies are useful, but they produce results which are difficult to interpret from the standpoint of how productivity contributes to societal goals. Economics is a study of the allocation of scarce resources to achieve competing ends. Production is not an end; consumption is. Thus, the productivity of an economy is best measured by the efficiency with which the economy provides consumption goods.

In our study, producers are considered together because their output is directly or indirectly necessary to produce goods which consumers use for similar purposes. Specifically, this is a study of the *food system* of the economy. We define food system as including exactly that portion of the activities of every industry involved in production and delivery of food for personal consumption. Following the productivity estimates, we discuss the contribution of selected industries to the system.

OUTPUT OF THE FOOD SYSTEM

The process by which the United States provides food for itself is obviously quite complex. Consequently, it may be suspected that any attempt to measure output, labor input, and, hence, labor productivity for the food system would be rather complicated. Whereas measurement of labor input is indeed difficult, our measurement of final output is straightforward. We used personal consumption expenditures for food as reported in the national income and product accounts, given in current and constant dollars, and disaggregated into food purchased for off-premises consumption, purchased meals and beverages, food furnished Government and commercial employees, and food produced and consumed on farms.

Table 1 shows the components of real personal consumption expenditures (PCE) for food in 1958, 1963, 1967, and 1970. (The choice of years is explained later.) Total PCE for food has grown at a rate approximately twice that of population growth. From 1958 to 1970, total U.S. population grew at an annual rate of 1.3 percent (*4*), whereas real PCE for food rose 2.5 percent annually.[1] All components of real PCE for food grew at about the same rate except for food produced and consumed on farms. The latter component declined at an average annual rate of 6.5 percent, partially reflecting the decline in farm population.

LABOR INPUT IN THE FOOD SYSTEM

Now we turn to a consideration of the measurement of labor input. All labor used within the entire economy to produce and distribute food must be identified. A complete measure includes not just the labor used by farmers, food processors, and distributors. It also considers the labor used in making inputs the three groups use directly and the labor used in making the inputs to produce the direct inputs.[2] It will be argued below that

---

*Eric C. Howe is a graduate student in economics at the University of Maryland. Gerald E. Schluter and Charles R. Handy are agricultural economists with the National Economic Analysis Division, ERS.

[1] Italicized numbers in parentheses refer to items in References at the end of this article.

[2] As a further complication, consider the problem of imports. The United States imports food for personal consumption and imports inputs used in producing food. The assumption was made that the United States pays for these imports with exports. Thus the labor involved in producing those exports had to be measured.

Table 1.—Personal consumption expenditures for food in constant (1958) dollars, selected years, 1958-70

Food use	1958[a]	1963[b]	1967[c]	1970[d]
	Million dollars[e]			
Food purchased for off-premises consumption	58,404	64,541	74,585	80,297
Purchased meals and beverages	15,319	17,613	19,693	20,265
Food furnished Government and commercial employees	1,243	1,169	1,684	1,694
Food produced and consumed on farms	1,410	825	623	630
Food for personal consumption	76,376	84,148	96,685	102,846

[a]U.S. Dept. Commerce. "Personal Consumption Expenditures in the 1958 Input-Output Study," *Survey of Current Business*. October 1965. [b]U.S. Dept. Commerce. "Personal Consumption Expenditures in the 1963 Input-Output Study,"*Survey of Current Business*, January 1971. [c]U.S. Dept.Commerce, "Input-Output Structure of the U.S. Economy: 1967," *Survey of Current Business*, February 1974. [d]U.S. Dept. Commerce. "U.S. National Income and Product Accounts," *Survey of Current Business*. July 1972. [e]The deflator for each component is the U.S. Department of Commerce implicit price deflator for that component as published in table 8.6 of the annual "National Income Issue" of the *Survey of Current Business*.

the only way to measure labor input this exhaustively is through the use of national input-output tables (I/O)[3] (For the interested reader, an algebraic statement of our methodology appears in the appendix.) The information available in the national I/O tables can be used to relate PCE for food to the labor needed in all parts of the economy to produce this food. We computed labor productivity for the recent years for which national I/O tables are available: 1958, 1963, and 1967.[4] For 1970 figures, the 1967 transactions matrix was used. The implications of re-use of the 1967 matrix are *not* intuitive and they are discussed after the algebraic model is developed in the appendix.

Total direct and indirect labor requirements appear in table 2. Direct labor requirement coefficients (the diagonal entries in the matrix L in the appendix) were computed as the number of civilian employees per unit of output. [5] Total labor required in the food system has remained relatively stable from 1958 to 1970. However, the labor force grew 1.66 percent annually, which implies that the share of the U.S. labor force needed to meet the food needs of the population is decreasing (*4*). Labor

Table 2.—Total direct and indirect labor requirements to produce food for personal consumption, selected years, 1958-70

Labor use	1958	1963	1967	1970
	Thousands of persons engaged			
Food purchased for off-premises consumption	9,480	8,805	8,900	9,077
Purchased meals and beverages	2,536	2,717	2,778	2,815
Food furnished Government and commercial employees	202	176	203	174
Food produced and consumed on farms	264	145	85	73
Food for personal consumption	12,482	11,844	11,965	12,139

---

[3] The idea of measuring productivity with the use of I/O is not new. One methodology has been examined theoretically by (*5*, pp. 79-108), and a review of the applications of I/O to productivity is contained in (*2*). Our specific methodology differs from all precursors.

[4] The 82-industry current transactions tables were used in each year. Since the current transactions tables do not include capital flows, the labor requirements shown in this article do not include the labor required in the capital goods industries (except labor necessary to produce spare parts for maintenance). A useful extension of the study would be to compute these additional labor requirements.

[5] The labor data are taken from worksheets on total civilian employment by industry, U.S. Dept. Labor, Bur. of Labor Statistics.

required to produce purchased meals and beverages has risen gradually and labor required for food produced and consumed on farms has declined markedly.

LABOR PRODUCTIVITY IN PRODUCING FOOD FOR PERSONAL CONSUMPTION

We measure labor productivity as the level of PCE divided by the number of people engaged in the food system (table 3). Note that labor productivity varies widely with the type of consumption and that this dispersion is increasing in magnitude. In 1958, highest value was $6,161 (for food for off-premises consumption) and its lowest, $5,341 (for food produced and consumed on farms); a range of 15 percent. By 1970, the range of this dispersion had grown to 35 percent: productivity was highest for food furnished Government and commercial employees, at $9,736, and lowest for purchased meals and beverages, at $7,199.

This growing dispersion reflects differing rates of productivity growth, which is evident from examining table 4. Growth in labor productivity for the production of food for personal consumption averaged 2.75 percent annually from 1958 to 1970. The growth rate was not constant, however; productivity grew rapidly during 1958-67 and slowed during 1967-70.[6] The results shown in table 4 verify conventional wisdom in several respects. The general slowdown in the rate of productivity advance during the late 1960's has been noted before, and it is reflected in the aggregate shown in the bottom line of table 4. Conventional wisdom further suggests that productivity advance would be lowest for purchased meals and beverages (because they are service-oriented), which is seen to be true. Also, the growth during 1958-70 in productivity is higher for food produced and consumed on farms than for food purchased for off-premises consumption. Because the former category consists of foods which require little transportation or marketing service, this change means, in essence, that labor productivity is

Table 3.—Output per person engaged in production of food for personal consumption, selected years, 1958-70

Type of food use	1958	1963	1967	1970
	1958 dollars			
Food purchased for off-premises consumption	6,161	7,330	8,380	8,846
Purchased meals and beverages	6,041	6,483	7,089	7,199
Food furnished Government and commercial employees	6,153	6,642	8,296	9,736
Food produced and consumed on farms	5,341	5,690	7,329	8,630
Food for personal consumption	6,119	7,105	8,081	8,472

Table 4.—Labor productivity, average annual growth rates

Food use	1958-1963	1963-1967	1967-1970	1958-1970
	Percent			
Food purchased for off-premises consumption	3.54	3.40	1.81	3.06
Purchased meals and beverages	1.42	2.26	.51	1.47
Food furnished Government and commercial employees	1.54	5.71	5.48	3.90
Food produced and consumed on farms	1.27	6.54	5.60	4.08
Food for personal consumption	3.03	3.27	1.59	2.75

---

[6] As noted above, the 1970 productivity estimate was made using the 1967 direct requirements matrix. It is unlikely that all slowdown in productivity during 1967-70 resulted from this re-use of the 1967 table—though some of it may have. Due to the way in which technological change impacts on an I/O matrix, there is no *a priori* reason to suppose that re-use of the 1967 direct requirements matrix understates growth in productivity. As noted in the appendix, it might easily be overstated. It will be possible to check whether the slowdown was real when the 1972 national I/O study becomes available.

growing faster for farming (and its input industries) than for marketing and transportation (and their input industries.)[7]

A number of provocative fluctuations in the rate of productivity advance can be seen by examining the three subperiods shown in table 4. However, relating these changes to specific technological innovations and to shifting input mixes, is beyond the scope of the study.

AN ECONOMY-WIDE PROCESS

In this section, we make our concept of the food system more concrete and support our assertion that only by using I/O can productivity be correctly measured.

The process by which the United States produces food for personal consumption is quite complex, as will now be established by showing the levels at which 28 selected industries participate in the food system.[8] We show participation in three ways. Table 5 presents gross output required from each of the 28 industries for the economy to produce food for personal consumption. Table 6 and 7 can both be related to table 5. Table 6 shows the labor needed in each of the 28 industries to produce the output appearing in table 5. Table 7 contains the income generated (or value added) in each of the selected industries.

---

[7] A word of caution is in order here. The I/O methodology assumes that the slaughtering of livestock consumed on farms is done in I/O industry 14 (food processing). Thus, some of the productivity advance shown for food produced and consumed on farms may be attributable to I/O 14.

[8] Generally, we selected the 28 industries whose participation was largest. However, some smaller industries were included when they seemed of particular interest.

Table 5.—Gross output needed from selected industries to produce food for personal consumption, selected years, 1958-70

Producing industry	1958	1963	1967	1970
	Million dollars			
1. Livestock and livestock products	21,465	20,931	24,522	30,249
2. Other agricultural products	13,406	15,620	17,490	19,968
3. Forestry and fishery products	581	723	858	1,027
4. Agricultural, forestry, and fishery services	961	1,037	1,376	1,829
7. Coal mining	362	292	299	501
8. Crude petroleum and natural gas	1,243	1,269	1,365	1,526
12. Maintenance and repair construction	1,902	1,713	2,061	2,665
14. Food and kindred products	59,036	64,981	78,011	94,005
24. Paper and allied products, except containers	1,933	2,756	3,181	3,712
25. Paperboard containers and boxes	1,203	1,534	2,054	2,391
26. Printing and publishing	1,684	2,127	2,611	3,110
27. Chemicals and selected chemical products	2,091	2,657	3,928	4,194
31. Petroleum refining and related industries	2,096	2,348	2,569	2,761
35. Glass and glass products	701	825	1,087	1,399
39. Metal containers	1,466	1,499	1,926	2,338
44. Farm machinery and equipment	220	240	310	346
59. Motor vehicles and equipment	789	735	823	1,019
65. Transportation and warehousing	8,422	7,734	9,336	11,133
66. Communications, except radio and TV broadcasting	915	1,089	1,460	1,646
68. Electric, gas, water, and sanitary services	1,927	2,445	2,862	3,278
69. Wholesale and retail trade	28,609	36,358	47,745	57,821
70. Finance and insurance	2,192	2,220	2,628	3,545
71. Real estate and rental	4,363	5,161	6,933	8,405
72. Hotels, personal and repair service, excluding auto	348	462	657	808
73. Business services	4,471	5,673	8,276	10,181
75. Automobile repair and services	877	840	1,193	1,504
77. Medical, educational services, and nonprofit organizations	276	300	376	509
80. Gross imports of goods and services	5,048	5,245	6,243	7,457
Other industries	14,455	15,876	21,290	25,555
Total	183,042	204,690	253,470	304,882

Table 6.—Direct and indirect labor requirements in selected industries to produce food for personal consumption selected years, 1958-70

Producing industry[a]	1958	1963	1967	1970
	Million dollars			
1.	2,019	1,652	1,390	1,176
2.	1,756	1,493	1,301	1,299
3.	23,	23,	26	26
4.	136	140	134	180
7.	29	17	14	17
8.	40	32	27	25
12.	97	98	103	81
14.	1,656	1,580	1,591	1,624
24.	73	90	88	90
25.	54	61	73	76
26.	128	133	133	140
27.	70	66	82	78
31.	26	20	17	17
35.	45	45	50	55
39.	51	44	45	46
44.	10	9	10	9
59.	20	14	15	17
65.	667	525	502	490
66.	76	59	64	64
68.	59	51	50	48
69.	4,069	4,323	4,721	4,905
70.	173	160	151	169
71.	52	48	50	54
72	67	78	89	98
73.	226	312	376	452
75.	42	36	43	45
77.	46	42	43	48
80	0	0	0	0
Other industries	772	693	777	810
Total	12,482	11,844	11,965	12,139

[a]For complete stub, refer to table 5.

Table 7.—Income generated in selected industries to produce food for personal consumption, selected years, 1958-70

Producing industry[a]	1958	1963	1967	1970[b]
	Million dollars			
1.	7,359	5,250	6,472	
2.	6,774	8,496	8,551	
3.	227	247	361	
4.	429	341	785	
7.	211	171	175	
8.	764	717	782	
12.	1,165	963	1,209	
14.	15,066	17,406	20,924	
24.	672	1,011	1,177	
25.	450	610	755	
26.	795	1,030	1,298	
27.	807	1,083	1,447	
31.	420	548	656	
35.	389	452	603	
39.	492	512	672	
44.	79	79	108	
59.	229	218	252	
65.	5,085	4,708	5,478	
66.	779	923	1,211	
68.	941	1,144	1,358	
69.	20,727	26,662	34,564	
70.	1,228	1,222	1,481	
71.	3,157	3,789	5,146	
72.	212	295	392	
73.	2,051	2,772	4,263	
75.	422	497	655	
77.	188	210	263	
80.	0	0	0	
Other industries	5,461	6,346	8,331	
Total	76,579	87,702	109,369	

[a]For complete stub, refer to table 5. [b]A conceptually consistent estimate of income generated cannot be computed for 1970.

Thus, each entry in table 7 is the output shown in table 5 minus purchases of intermediate products.[9]

These three tables present some interesting numbers. Note that the farming industries' participation in the food system is dwarfed by that of the other industries. For example, in 1967, production of food for personal consumption generated $253 billion of total business activity but only $42 billion (17 percent) represents output from the agricultural industries (I/O industries 1 and 2). Similarly, the agricultural industries retain only 14 percent of the income generated. One implication of these results is that the price which consumers pay for food is relatively insensitive to factor compensation (returns to labor and capital) in agriculture.[10] The livestock and livestock products industry required 2.02 million workers to produce food for personal consumption in 1958. By 1970, its labor requirements had fallen to 1.18 million workers. The proportion of total labor requirements from all industries accounted for by the two agricultural industries fell from 30.3 percent in 1958 to 20.4 percent in 1970.

---

[9] More productivity measures would result from deflating each entry in table 7 and dividing by the corresponding entry in table 6. This was not done, for lack of appropriate price deflators. In addition, the measure would factor to a ratio of the direct employment requirements per dollar of real income retained, ignoring indirect employment requirements, and thus not conforming to the approach to productivity taken in this article. Using these additional productivity numbers, we might seemingly measure the contribution of each industry to the overall increase in productivity shown above, but the theoretical basis for such an analysis would have to be developed. At best, the numbers alone would have little significance.

[10] From 1972 to 1973, factor compensation in farming rose 60 percent while the implicit price deflator for PCE for food rose 13 percent, a ratio consistent with our results (*3*).

While food processing (I/O 14) contributed 31 percent of total business activity required to produce food for personal consumption in 1967, its proportion of total value added, 19 percent, and of total labor requirements, 13 percent, were much smaller.

Conversely, wholesale and retail trade has the highest labor requirements (40 percent of the total in 1967) and the highest proportion of value added (32 percent). But the proportion of total business activity is much smaller (19 percent).

Both the trade and service industries are expanding in importance within the food system. Wholesale and retail trade increased its share of total gross output from 15.6 percent in 1958 to 19.0 percent in 1970, and its share of labor requirements from 32.6 percent to 40.4 percent. Business services (I/O 73) increased its share of gross output required to produce food for PCE from 1.1 percent to 3.3 percent from 1958 to 1970.

It is commonly realized that expenditures for chemicals (I/O 27) have been increasing due to more outlays for fertilizers, pesticides, and food additives. Indeed, the nominal dollar output of I/O 27 (which is used to produce food for personal consumption) has been growing at an average annual rate of 5.9 percent. Yet, this rapid increase has only raised I/O 27's share of total business activity from a 1958 value of 1.14 percent to a 1970 value of 1.37 percent.

I/O data must be used to measure labor input exhaustively because production of food for domestic personal consumption is an economy-wide process. Farming, food processing, transportation, wholesaling, and retailing certainly predominate, yet in 1970 they accounted for less than four-fifths of total labor input. Exhaustively measuring labor input requires the identification of the sector location of all economic activity supporting the production of food, which necessitates the use of I/O.

APPENDIX

The fundamental identity for input-output analysis is:

$$X = A X + Y \qquad (1)$$

where A is an *nxn* matrix of direct requirement coefficients, X is an *nxl* vector of outputs, and Y is an *nxl* vector of final demands.

For the study, A was taken from the 82-industry national I/O model, as noted. By taking Y, then, as the vector of final demands for food for personal consumption (available in the sources listed in the footnotes to table 1), equation (1) could be used to compute output requirements, X, for the production of food for personal consumption. Labor requirements could be easily computed by multiplying each industry's output requirements by its labor coefficient (labor per unit of output).[11] The problem with this approach is that it treats imports as though they were free. It ascribes to domestic labor some part of the productivity of foreign workers. Since imported goods are not free, their cost must be accounted for.

Imports must be considered in two contexts with respect to the way in which the U.S. economy provides food for its own personal consumption. Some food is directly imported. In addition to the (relatively small) direct imports of food, almost every industry in the economy (at the 82-industry level of aggregation) uses some imports directly in its production processes. As pointed out earlier, we assume that the United States pays for its imports of goods by exporting goods.[12] Thus, when output requirements are computed for the production of food for personal consumption, they must be adjusted upward to include output which was used in the production of those exports which were necessary to pay for imports. An algebraic statement of our methodology follows.

Equation (1) must be altered to make exports endogenous. As a first step, include exports in equation (1) which results in

$$X = A X + Y + E e \qquad (2)$$

where E is an *nxl* vector giving the average share of each industry in $1 of exports and *e* is an *lxl* matrix of exports. Industry 80 in the 82-industry matrix is an aggregate import industry. If M denotes an *lxn* vector containing the 80th row of $(I - A)^{-1}$, total import requirements will be M(Y + E*e*). Imposition of the constraint that imports are paid for with exports results in the condition:

$$e = M(Y + E e) \qquad (3)$$

which implies

$$e = \frac{1}{1 - M E} M Y \qquad (4)$$

Substituting equation (4) into equation (2) and solving for X results in

$$X = (I - A)^{-1} \left(I + E \frac{1}{1 - M E} M\right) Y \qquad (5)$$

Equation (5) may be simplified by noting that

$$\left(I + E \frac{1}{1 - M E} M\right) = (I - EM)^{-1} \qquad (6)$$

---

[11] For an illustration of this approach, see (*1*).

[12] This is a perfectly reasonable assumption *in the long run*. It does however, ignore the fact that the U.S. *annual* balance of trade could be negative. In that case, the United States would be paying for imports with I O U's. Since the balance of trade was positive in 1958, 1963, 1967, and 1970, the assumption is justified.

and, substituting equation (6)[13] into (5) and simplifying

$$X = (I - A - \widetilde{E})^{-1} Y \qquad (7)$$

where $\widetilde{E}$ is an *nxn* matrix which has the E vector in its import column and zeroes elsewhere.

Let L and V denote two *nxn* diagonal matrices of labor coefficients and value added coefficients, respectively. Then labor requirements are given by the expression L X and income generated by the expression V X.

Our measure of output per person engaged is:

$$\frac{u\, Y\, P^{-1}}{u\, L(I - A - \widetilde{E})^{-1}\, Y} \qquad (8)$$

where u is a row vector of units and P^{-1} is the reciprocal of the implicit price index for PCE for food. Equation (8) may be altered to a more useful form. Let $f = u\, Y$ be the total nominal dollar PCE for food and let $K = Y \frac{1}{f}$ be an *nxl* vector showing each industry's direct share of a dollar expenditure for food for personal consumption. By substitution and simplification, labor productivity is:

$$\frac{1}{P\, u\, L(I - A - \widetilde{E})^{-1}\, K} \qquad (9)$$

These steps indicate clearly that labor productivity depends computationally on the composition of final demand (K), the labor coefficients (L), the direct requirements coefficients (A), the composition of foreign demand for U.S. exports ($\widetilde{E}$), and the price index (P).

As stated earlier, we re-used the 1967 A matrix for the 1970 computations though adjusting it to reflect changes in relative prices between 1970 and 1967. For 1970, the matrices L, P, and E were computed as for other years. The composition of final demand was kept constant *within* each of the four components of PCE for food, but K was varied due to the change in the relative weight of each component. It will be possible to check whether this procedure was justified when the 1972 national I/O study is published.

It might seem that this procedure for 1970 must necessarily understate productivity growth. That is not true. An industry may save on labor by the use of more purchased inputs, thus increasing some elements of A but decreasing an element of L. Therefore, the bias could be either way. Intuitively, it seems that the procedure is justified if $L(I - A - \widetilde{E})^{-1} \widetilde{K}$ has approximately the same Frobenius eigenvalue with the 1967 A as with the 1970 A, but that assertion has not been proven. ($\widetilde{K}$ is a diagonal matrix containing the elements of K.)

REFERENCES

(1) Gale, Hazen F. "Industry Output, Labor Input, Value Added, and Productivity." *Agr. Econ. Res.* Vol. 20, No. 4, October 1968, pp. 113-133.

(2) Howe, E. C. and Handy, C. R. "Inventory and Critique of Productivity Estimates in the U.S. Food and Fiber Sector," *Amer. J. Agr. Econ.* Vol. 57, No. 5, December 1975, pp. 916-921.

(3) U.S. Department of Commerce, Bureau of Economic Analysis. *Survey of Current Business.* July 1974.

(4) U.S. Department of Commerce. *Statistical Abstract of the United States.*

(5) Rymes, T. K. "The Measurement of Capital and Total Factor Productivity." *Rev. Income and Wealth,* Vol. 18, 1972.

---

[13] Equation (6) may be proven by noting that, by a power expansion and simplification:

$$\begin{aligned}(I - E\,M)^{-1} &= I + E\,M + (E\,M)^2 + (E\,M)^3 + \dots \\ &= I + E(1 + M\,E + (M\,E)^2 + (M\,E)^3 + \dots)M \\ &= I + E\,\frac{1}{1 - M\,E}\,M\end{aligned}$$

PRODUCTIVITY CHANGES IN THE FOOD AND FIBER SYSTEM, 1958-74

By Donald D. Durost and James E. Kirkley*

Measures of output, input, and labor productivity are derived for the food and fiber sector. They are used to analyze changes in output, hours worked, and labor productivity that have occurred in the food and fiber system since 1958. A major result of the analysis is that the reduction in labor at the farm level has apparently not been reflected in either the input or product marketing sectors.

Keywords: Food and fiber sector, output, hours worked, partial productivity.

The U.S. food and fiber system consists of several industries responsible for producing both agricultural goods and the inputs needed in production of these goods. Food and fiber industries are also responsible for maintaining flows of products and necessary inputs from the basic resource stage to the final consumer. To move these goods and inputs requires considerable interaction among the industries within the system's four sectors: inputs, farm, product market, and consumer. The first three sectors produce the products and inputs and move them through to the fourth sector. The consuming sector creates the demand that generates the flow.

In this article, we focus primarily on the three producing sectors and on effects of major changes within these sectors in recent decades. These changes have helped transform U.S. food production into a highly sophisticated, complex operation.

Many of these changes have caused the scale and size of off-farm operations to expand. The input sector has grown as farmers have become more dependent on purchased or nonfarm inputs. In 1947, such inputs made up approximately 44 percent of all inputs used on the farm. Currently, they account for over 60 percent.

In the farm sector, farming activities have become increasingly complex, and many of the jobs performed have been modified. Farmers have had to adapt to the use of modern inputs and comply with changing market requirements. At one time, farmers produced the type, size, and quality of products that they thought best and ran the risk of finding that the market wanted something different. More and more, they are learning the value of paying closer attention to what the market demands. Many agricultural commodities are now produced according to buyers' specifications, for example, formula fed broilers of a specific age and weight, cattle of an exact weight and finish, or wheat with a minimum level of protein. Product specification is a principal feature of production contracts between farmers and buyer-processors.

The product market sector, once comprised of small local markets participated in by nearby producers and processors, and consumers, has changed significantly. The development of new processing technologies and products has promoted large processing and distribution operations. Often, the processing plant is associated with the retail outlet. These large operations have encouraged standardization of products, which, in turn, has forced standardization of commodities produced on the farm.

Increasingly, consumers, the fourth sector in the system, can buy more and more convenience foods. These items have transferred the preparation time, culinary skills, and energy inputs from homemakers' kitchens to the food processor and distributor. In 1973 alone, food companies introduced an estimated 7,200 new products or variations of these for the retail market.

The rapid changes occurring in the food and fiber producing sectors are bound to have some unsettling effects. These have the potential of creating problems for policymakers. Though they have some information, policymakers lack adequate data about changes in output, input, and productivity.

This article attempts to fill the void by presenting measures of output, input, and productivity and possible answers to the following questions:

- What has happened to output?
- Has labor reduction at the farm level been transferred to the other sectors?
- What has happened to labor productivity?

MEASURING OUTPUT

Output may be defined in two ways: (1) gross output minus the value of products produced and used within each sector for further production, and (2) net output, which is the value added by each sector. Both output and input must be measured in constant dollars so that dissimilar items can be combined; for example, corn and cotton, labor and capital.

The net output approach can be used to determine what product or products can be attributed to the inputs within a sector. It differs from the gross output approach in that it excludes intermediate products and results in different measures of input shares. However, the method is useful for measuring the contribution of a subsector to a total sector or system.

---

*Donald D. Durost and James E. Kirkley are agricultural economists in the National Economic Analysis Division, ERS.

The gross output approach can be used to measure the volume of output of a subsector. However, these volumes cannot be added to get a sector's total as can be done with the net output approach. But because the U.S. Department of Agriculture is primarily interested in the volume of output from all resources committed to food and fiber, the gross output concept is used in this article.

Gross output of the food and fiber system is defined to include the constant dollar value of the following: (1) food expenditures by civilians, (2) value of goods produced on the farm for home consumption, (3) agricultural exports, and (4) manufactured values of tobacco products, beverages, leather tanning, and weaving and finishing of cotton and wool.

The constant dollar values of various items are derived by deflating current values with appropriate price indexes as follows: food expenditures and home consumption values with the consumer food index, exports with the wholesale farm products and processed food index, and manufactured products with individual wholesale prices for census years. However, the constant dollar value for noncensus years had to be interpolated (table 1).

Two problems arise with the gross output measure: use of manufactured values as a proxy for consumer purchases for some products, and omission of certain goods.

The first problem is not too serious because manufactured values should provide a relatively good indication of consumer purchases of the products in category four. However, the use of manufactured values as a proxy for consumer purchases could possibly result in an error in measuring output. Omission of certain goods may be a more severe problem because the omission could result in an understatement of productivity. Several items are excluded from the constant dollar value of output because data are not readily available for the following items: the share of value accounted for by natural fibers as opposed to manmade fibers in blended products,[1] expenditures for imported foods and all fish, domestic military food and fiber purchases, inventory changes, nursery and flower products, and markets for oils (soaps, paints, and so on).

MEASURING INPUTS

Sound measures of inputs are obviously needed to determine what each input contributes to output. However, the U.S. food and fiber system uses several intermediate products as inputs for processing, producing, and marketing final output, which poses a difficulty.

Table 1.—Value of food and fiber output and percentage distribution, 1958-74
(1967 dollars)

Year	Food (domestic use)[a]		Exported		Other products (domestic use)[b]		Total	
	Value	Share	Value	Share	Value	Share	Value	Share
	Mil. dol.	*Pct.*	*Mil. dol.*	*Pct.*	*Mil. dol.*	*Pct.*	*Mil. dol.*	*Pct.*
1958	70,620	78.4	3,930	4.4	15,488	17.2	90,038	100.0
1959	74,521	78.7	4,230	4.4	15,989	16.9	94,740	100.0
1960	77,370	78.1	5,157	5.2	16,489	16.7	99,016	100.0
1961	78,319	77.8	5,362	5.3	16,989	16.9	100,670	100.0
1962	80,435	77.9	5,316	5.2	17,489	16.9	103,240	100.0
1963	82,198	77.5	5,953	5.6	17,990	16.9	106,141	100.0
1964	84,781	77.0	6,811	6.2	18,529	16.8	110,131	100.0
1965	86,785	77.3	6,415	5.7	19,089	17.0	112,189	100.0
1966	88,544	77.1	6,648	5.8	19,639	17.1	114,831	100.0
1967	89,965	77.2	6,380	5.5	20,190	17.3	116,535	100.0
1968	91,410	77.4	6.082	5.1	20,653	17.5	118,145	100.0
1969	91,444	77.5	5,496	4.6	21,116	17.9	118,056	100.0
1970	92,840	76.8	6,499	5.4	21,579	17.8	120,918	100.0
1971	94,103	76.6	6,760	5.5	22,042	17.9	122,905	100.0
1972	95,086	75.9	7,681	6.1	22,503	18.0	125,270	100.0
1973	94,101	73.3	11,113	8.7	23,105	18.0	128,319	100.0
1974	92,063	71.8	12,398	9.7	23,705	18.5	128,166	100.0

[a] Includes food expenditures and home consumption on farms. [b] Includes manufactured values for census years 1958, 1963, 1967, and 1972, and interpolations for other years. Includes tobacco products (SIC 21), beverages (SIC 208), leather tanning (SIC 311), textiles (SIC 221, 223, 2261).

---

[1] Because these proportions could not be determined, only products made of natural fibers were included.

Unfortunately, for most such inputs, data either are not readily available or are very limited in coverage. Therefore, only labor, in terms of hours worked (man-hours) were used in the study.

Data on hours worked came from the following sources: (1) for farming, from the U.S. Department of Commerce's Bureau of Labor Statistics, (2) for marketing and processing farm foods, from the Economic Research Service, and (3) for tobacco, beverages, leather, and textiles, from the various censuses of manufacturing.

Since data on hours worked were incomplete for the farm input sector, estimates were derived with a method developed by Folke Dovring (*2*).[2] His procedure included farm expenditures for variable inputs plus capital depreciation. He adjusted this sum to exclude the farm value of feed, seed, and livestock purchases. Labor's share of inputs and services was obtained by assuming the same labor ratio as that of total national wages to national income. Hours worked were obtained by dividing the labor share of the value of inputs by the wage rates in the manufacturing subsector (table 2). This method inherently implies that all labor is accounted for, from the raw materials stage to delivery at the farm. But there are some major weaknesses in determining hours worked. First, the share of national income from employee compensation can change from year to year because of fluctuations in general economic conditions. Second, differential rates of inflation among farm input expenditures and the effects of inflation on real wage rates can bias the hours worked. However, the relative changes or trend in hours worked, as derived by this method, probably gives an accurate description of the changes in hours worked (table 3).

MEASURING PRODUCTIVITY

Productivity represents the quantity of output obtainable from a given set of inputs during a given period of time. There are, in general, two types of productivity measurements: an average or gross measure where all inputs are considered, and a measure

Table 2.—Computation of hours worked in farm input sector, 1958-74

Year	National income[a]	Compensation to labor[b]	Labor share of national income[c]	Farm input expenditures		Manufacturing wage rate[f]	Hours worked[g]
				Total value[d]	Labor share[e]		
	Bil. dol.		*Pct.*	*Bil. dol.*		*Dol.*	*Mil. h.*
1958	361.2	255.4	70.71	13,776	9,741	2.11	4,617
1959	394.2	276.7	70.19	14,800	10,388	2.19	4,743
1960	409.5	291.8	71.26	14,797	10,544	2.26	4,665
1961	421.0	300.4	71.35	15,146	10.807	2.32	4,658
1962	454.1	321.8	70.87	15,791	11,191	1.39	4,682
1963	479.4	339.5	70.82	16,441	11,644	2.46	4,733
1964	515.7	364.5	70.68	16,823	11,890	2.53	4,700
1965	562.4	392.9	69.86	17,431	12,177	2.61	4,666
1966	618.5	435.6	70.43	18,681	13,157	2.72	4,837
1967	652.1	468.2	71.80	19,976	14,343	2.83	5,068
1968	710.5	515.9	72.61	20,528	14,905	3.01	4,952
1969	763.7	567.2	74.27	21,429	15,915	3.19	4,989
1970	794.1	604.9	76.17	23,363	17,796	3.36	5,296
1971	853.7	654.9	75.66	24,015	18.120	3.57	5,090
1972	947.4	710.5	75.00	25,359	19,019	3.81	4,992
1973	1,062.1	792.5	74.62	30,389	22,676	4.08	5,558
1974	1,135.1	867.0	76.38	37,185	28,402	4.41	6,440

[a] Excludes farm wages. National income from 1976 *Economic Report of the President*, Table B-12. [b] Excludes farm wages. Data from the sources in footnote a. [c] Compensation to labor divided by national income. [d] Expenditures include: fertilizer and lime; repair and operations; feed, seed, and livestock purchases (excluding farm value); depreciation and capital consumption; and miscellaneous. Farm operator's dwellings and non-real estate interest excluded. Data from various tables from *Farm Income Statistics*, Stat. Bul. 547, Econ. Res. Serv., USDA. [e] Value of farm input expenditures times labor's share of national income. [f] Manufacturing wage rates from the 1976 *Economic Report of the President*, Table B-18. [g] Wage rates divided into labor's share of farm input expenditures.

---

[2] Italicized numbers in parentheses refer to items in References at the end of this article.

Table 3.—Hours worked in the food and fiber system, 1958-74

Year	Farm input market		Farm production		Product market		Total	
	Mil. h.[a]	*Pct.*	*Mil. h.*[b]	*Pct.*	*Mil. h.*[c]	*Pct.*	*Mil. h.*	*Pct.*
1958	4,617	16.0	12,906	44.6	11,375	39.4	28,898	100.0
1959	4,343	15.1	12,919	45.0	11,481	39.9	28,743	100.0
1960	4,665	16.4	12,463	43.7	11,385	39.9	28,513	100.0
1961	4,658	17.0	11,796	43.0	10,981	40.0	27,435	100.0
1962	4,682	17.4	11,434	42.3	10,889	40.3	37,005	100.0
1963	4,733	17.9	10,955	41.4	10,745	40.7	26,433	100.0
1964	4,700	18.2	10,359	40.0	10,823	41.8	25,882	100.0
1965	4,666	18.1	10,091	39.2	11,009	42.7	25,766	100.0
1966	4,837	19.3	9,141	36.5	11,071	44.2	25,049	100.0
1967	5,068	20.3	8,815	35.3	11,096	44.4	24,979	100.0
1968	4,962	19.8	8,624	34.5	11,167	44.7	24,753	100.0
1969	4,989	20.4	8,152	33.3	11,307	46.3	24,448	100.0
1970	5,296	22.0	7,645	31.7	11,169	46.3	24,110	100.0
1971	5,090	21.4	7,423	31.2	11,258	47.4	23,771	100.0
1972	4,992	20.9	7,475	31.3	11,444	47.8	23,911	100.0
1973	5,558	22.8	7,342	30.1	11,491	47.1	24,391	100.0
1974	6,440	25.1	7,352	28.7	11,859	46.2	25,651	100.0

[a] See table 2. [b] Data from Bureau of Labor Statistics, U.S. Dept. Commerce. [c] Data for processing and marketing of farm food products from *Marketing and Transportation Situation*, MTS-198, Econ. Res. Serv., USDA, table 11. Data for nonfood manufactured products are from Industrial Censuses for 1958, 1963, 1967, and 1972, and they are interpolated for other years.

which considers only individual or subsets of inputs.

The gross measure of productivity is difficult to construct because data are not adequate for many of the inputs.[3] Thus, the study involved only a measure of productivity of a single input; namely, labor. With other input changes held constant, this measure credits all changes in output for the system as a whole to labor (table 4). However, as will be shown later, different rates of change in labor productivity are implied for the three sectors by differences in the way each of them has apparently combined other inputs with labor.

CHANGES IN OUTPUT AND LABOR SINCE 1958

In the overview section, three questions were asked:

- What has happened to output?
- Has the reduction in labor at the farm level been transferred to the other sectors?
- What has happened to labor productivity?

Food and fiber output has changed considerably since 1958 (table 1). Total output (constant dollars) increased from $90,038,000 that year to $128,166,000 in 1974, a

Table 4.—Food and fiber labor productivity indexes, 1958-74
(1967 = 100)

Year	Output[a]	Hours worked[b]	Labor productivity[c]
1958	77.3	115.7	66.8
1959	81.3	115.1	70.6
1960	85.0	114.2	74.4
1961	86.4	109.8	78.7
1962	88.6	108.1	82.0
1963	91.1	105.8	86.1
1964	94.5	103.6	91.2
1965	96.4	103.2	93.4
1966	98.5	100.3	98.2
1967	100.0	100.0	100.0
1968	101.4	99.1	102.3
1969	101.3	97.9	103.5
1970	103.8	96.5	107.6
1971	105.5	95.2	110.8
1972	107.5	95.7	112.3
1973	110.1	97.7	112.7
1974	110.0	102.3	107.5

[a] Index based on constant values shown in table 1. [b] Index based on hours worked shown in table 3. [c] Derived by dividing index of output by index of hours worked.

---

[3] The only regularly available measure of gross productivity is for the farm production sector. This measure is prepared by the National Economic Analysis Division, Economic Research Service.

42-percent rise. Output went up every year except 1974, when no change occurred from the 1973 level.

Food has accounted for the greatest share of total food and fiber output, ranging from 78.7 percent in 1959 to 71.8 percent in 1974. But this share has declined slightly and persistently. Further, estimates of food's share may be on the high side because of certain omissions from total output previously mentioned.

Thus, food's share of total output decreased during the period although total volume of food has increased. It has gone from $70,620,000 in 1958 to $92,063,000 in 1974—a steady increase in all years except 1973 and 1974. Possibly, the slight decrease between 1973 and 1974 occurred because of rising food prices and declining per capita consumption in those years.

Exports have more than tripled since 1958; however, most of this increase occurred between 1972 and 1974. Exports' share of total food and fiber output has remained relatively constant except for the last 3 years. In that period, the share rose to almost 10 percent (1974).

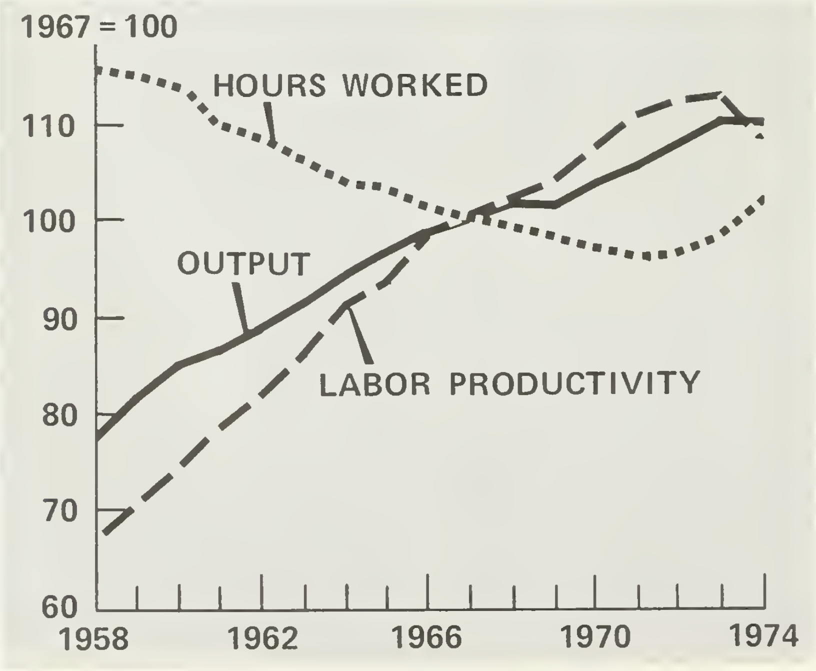
For labor's input though, the story is quite different. Since 1958, total number of hours worked fell continuously until 1972, with all of the decrease accounted for by the farm production sector (table 3). The recent modest decline for farm labor, however, has been more than offset by increases in the other two sectors.

In fact, the upward swing in total hours worked in 1973 and 1974 came largely because of the sharp increase in farm input labor. Part of this increase is real because farmers have purchased record levels of inputs. However, part of the rise results from the method used to calculate the hours. The inflation rate for farm inputs appears to have outstripped the real wage rates for manufacturing. Up until 1973, however, hours worked in the farm input sector moved up very slowly.

The product market sector has shown remarkable stability in the use of labor over the entire 17-year period. Use has remained at just over 11 billion hours, except for a slight drop in the early 1960's and a slight increase in 1974. For farm production, hours worked have fallen each year since 1959 except in 1972 and 1974.

These different trends indicate that labor-saving inputs and technologies increased output per hour worked in farming, but reduced the number of hours needed. In the other two sectors, input substitution in-

FOOD AND FIBER OUTPUT, HOURS WORKED, AND PRODUCTIVITY, 1958-74



USDA NEG. ERS 2559-76 (9)

creased the productivity of labor but failed to reduce the number of hours.

In 1958, agriculture accounted for 45 percent of the hours worked in the food and fiber system. This share has declined each succeeding year. In 1974, slightly less than 30 percent of the hours worked were devoted to farming. It has been theorized that the reduction in labor required for farming has been picked up in other nonfarming occupations, including the input and product market sectors. The evidence suggests that this shift has occurred only for occupations outside the food and fiber system.

Although hours of labor required decreased 14 percent in the food and fiber system between 1958 and 1974, the volume of food and fiber output increased 42 percent. Thus, labor productivity has risen 61 percent, or at an annual rate of 3.2 percent, since 1958 (table 4).

The greatest increase in labor productivity—47 percent—occurred during the first half of the period (figure). After 1966, it went up 15 percent. During the entire period, labor productivity decreased only once, a 5-percent drop which occurred between 1973 and 1974. Output had a similar trend but it experienced decreases in two periods—1968-69 and 1973-74. The hours worked, however, had a reverse trend. They consistently decreased until 1972, when they began to increase.

Although output has been increasing, not all of this rise can actually be attributed to labor. As suggested before, new technology directed at reducing labor and increasing productivity has been adopted by all three sectors. But its impact has been the most significant in farming.

CONCLUSIONS

All of the desired data are not available, but the measures of output, input, and productivity do provide information about the general direction of change in the U.S. food and fiber system. It does appear that labor productivity has been increasing. Furthermore, the technologies adopted by the farming sector have reduced the quantity of labor used and increased its productivity. Technologies adopted in the other two sectors, though, seem to have done more to increase productivity than to reduce labor requirements.

REFERENCES

(1) Arthur, H. B., Goldberg, R. A., and K. M. Bird. *The United States Food and Fiber System in a Changing World Environment.* Technical papers prepared for National Advisory Commission on Food and Fiber, August 1967.

(2) Dovring, Folke. *Productivity of Labor in Agricultural Production.* Univ. Ill. Agr. Exp. Sta. Bul. 726, September 1967.

(3) Hardin, Charles M. *Food and Fiber in the Nation's Politics.* Report prepared for the National Advisory Commission on Food and Fiber, August 1967.

(4) National Advisory Commission on Food and Fiber. *Food and Fiber for the Future, Report of the National Advisory Commission on Food and Fiber.* August 1967.

LEAST SQUARES ROBUSTNESS WHEN THE ERROR TERM IS MISSPECIFIED IN COBB-DOUGLAS TYPE FUNCTIONS

By Terry N. Barr and James F. Horrell*

The Cobb-Douglas type function has a multiplicative deterministic form, $Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2}, \ldots, X_k^{\beta_k}$, which can be used to analyze certain economic structures. The usual error specification for this function is that the error term enters multiplicatively with lognormal distribution. This specification permits use of a natural logarithmic transformation and the usual linear least squares estimation techniques with slight modifications. The study reported on investigates the effect of an error misspecification on the estimates of the parameters. In particular, is there a significant bias in the estimates of the parameters $\beta_1, \beta_2, \ldots \beta_k$ when the logarithmic transformation is applied to data generated using additive errors? Monte Carlo techniques were used to generate empirical distributions of the estimates to be analyzed with analysis of variance (ANOVA) techniques. Evidence suggests bias would exist but not of the magnitude, in most cases, to greatly challenge previous economic evaluations based on regressions using Cobb-Douglas type functions. Though this may be a source of consolation for many analysts, it should be noted that any conclusions that loglinear estimation yields satisfactory estimates in the presence of additive errors must still depend on the peculiarities of the sampling experiment.
Keywords: Production functions, econometrics, error misspecification.

INTRODUCTION

Economists have traditionally used the following Cobb-Douglas type functions in studies of demand and production (*4*, pp. 464-72).[1]

$$Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \ldots X_n^{\beta_n} \tag{1}$$

The error specification typically facilitates the use of the logarithmic transformation and the usual least squares estimation techniques with slight modifications (*6*, pp. 1034-38). The multiplicative error assumption can be rationalized through the multiplicative central limit theroem. But its pragmatic justification, namely allowing the Cobb-Douglas form to be intrinsically linear, has been eroded with the advance of nonlinear estimation techniques (*2*, pp. 101-102).

The question for the researcher is how serious is the bias in estimation if the assumption of the multiplicative error is incorrect and additive errors are present? Goldfeld and Quandt (*5*, pp. 251-57) have proposed a method of estimation devised to account simultaneously for additive and multiplicative errors. Although the technique appeared manageable and the results reasonable, a number of interesting statistical questions remained open. In particular is the question of the need to know more about the consequences of misspecification of the error terms. The problem is also discussed in many econometrics texts (*3*, pp. 213-18) (*11*, pp. 591-94). This article is designed to provide some insights into this question.

Two error specifications which are the most prevalent in the literature are:

$$Y_i = \beta_0 X_i^{\beta_1} e^{u_i} \qquad \text{U is N}(0, \sigma^2) \tag{2}$$

$$Y_i = \beta_0 X_i^{\beta_1} W_i \qquad \text{W is lognormal with mean } e^{\frac{1}{2}\sigma_v 2} \tag{3}$$

Suppose the *true model* which generated the Y_i's is

$$Y_i = \beta_0 X_i^{\beta_1} + V_i \qquad V_i \text{ is N}(0, \sigma^2 I) \tag{4}$$

But the specification used for estimation is the traditional model of (2). The conclusions will apply equally well to (3) but notational convenience dictates the alternative form. Employing the natural logarithmic transformation, equation (2) becomes

$$\ell n\, Y_i = \ell n \beta_0 + \beta_1 \ell n X_i + u_i \tag{5}$$

Upon examination it is clear that (5) is a linear model and the OLS estimates of $\ell n \beta_0$ and β_1 are given by

$$\ell n b_0 = \frac{1}{n} \sum_{i=1}^{n} \ell n Y_i - b_1 \frac{1}{n} \sum_{i=1}^{n} \ell n X_i = \overline{\ell n Y} - b_1 \overline{\ell n X} \tag{6}$$

---

*Terry N. Barr is an economist with the National Economic Analysis Division, ERS, and James F. Horrell is an associate professor of business administration at the University of Oklahoma.

[1] Italicized numbers in parentheses refer to items in References at the end of this article.

$$b_1 = \sum_{i=1}^{n} (\ell X_i - \overline{\ell n X})\ (\ell n Y_i - \overline{\ell n Y}) \Big/ \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})^2 \qquad (7)$$

By noting that

$$\sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})k = k \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X}) = k \cdot 0 = 0 \qquad (8)$$

we have, upon simplifying (7)

$$b_1 = \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})\ell n Y_i \Big/ \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})^2 \qquad (9)$$

Note from (4) that

$$\ell n Y_i = \ell n \beta_0 X_i^{\beta_1} (1 + V_i / \beta_0 X_i^{\beta_1}) = \ell n \beta_0 + \beta_1 l_n X_i + \ell n (1 + V_i / \beta_0 X_i^{\beta_1}). \qquad (10)$$

Substituting (10) into (9) and using (8) yields

$$b_1 = \left[\beta_1 \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})\ \ell n X_i\right] / \mathrm{D} + \left[\sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})\ \ell n (1 + V_i / \beta_0 X_i^{\beta_1})\right] / \mathrm{D}$$

where $D = \sum_{i=1}^{n} (\ell n X_i - \ell n X)^2$. The coefficient on β_1 reduces to 1, thus

$$E(b_1) = \beta_1 + \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X}) E\left[\ell n (1 + V_i / \beta_0 X_i^{\beta_1})\right] \Big/ \sum_{i=1}^{n} (\ell n X_i - \overline{\ell n X})^2 \qquad (11)$$

Looking at the term on the right-hand side of (11) that has the expected value operator, E, applied to it, we note the following. *Analytically* it is impossible to assume a multiplicative error term and make a logarithmic transformation when the error term is actually additive and normally distributed unless the error term, when negative, does not dominate the deterministic part of the model. For analytical purposes, it is necessary to truncate the error term distribution so that taking the logarithm of negative numbers can be avoided. Thus to obtain the expected value of b_1, it is necessary to evaluate

$$E\left[\ell n (1 + V_i / \beta_0 X_i^{\beta_1})\right] = K \int_{-\beta_0 X_i \beta_1}^{\infty} l_n (1 + V_i / \beta_0 X_i^{\beta_1}) \exp\left[-1/2 \frac{(V_i - \mu)^2}{\sigma}\right] dV_i$$

This evaluation will be carried out empirically with the use of Monte Carlo and ANOVA techniques. The Monte Carlo technique will be employed to generate empirical distributions of the estimates in accord with two specific stochastic functional forms.

Functional Form I: $Y_i = \beta_0 X_i^{\beta_1} + V_i$

Functional Form II: $Y_i = \beta_0 X_{1i}^{\beta_1} X_{2i}^{\beta_2} + V_i$

Once construction of the distributions is complete, analysis of variance (ANOVA) techniques employing factorial models can be utilized to analyze the effects on parameter bias of varying the sample characteristics.

DATA GENERATION AND MONTE CARLO APPLICATION

To illustrate the generation of the sample data for regression analysis, consider the single-variable model

$$Y_i = \beta_0 X_i^{\beta_1} + V_i$$

The value of the deterministic part of this model, namely $\beta_0 X_i^{\beta_1}$, is generated for each possible combination of β value, X range, and sample size identified in table 1. For all combinations the value of β_0 remains at 4.0.

For example, utilizing the first variant of each characteristic, a set of Y's would be constructed for $\beta_0 = 4$, $\beta_1 = .5$ in which X_i takes on values uniformly in the interval [5, 15] with a sample size of twenty, ($i = 1, \ldots, 20$). All of the generated (X, Y) points would lie on the graph of the function $Y = 4X^{.5}$. For example, if X = 5, the nonstochastic Y = 8.94. The Monte Carlo technique is used to obtain the probability distribution of estimators of any parameter for comparison with the true value of the parameter. By repeatedly adding random error values V_i from a known distribution to the deterministic portion of each combination and performing regression analysis, we derive an empirical distribution of each estimator.

Table 1.—Sample characteristics for functional form I

β values	X range	Sample size	Error variance
1. $\beta_1 = .5$	1. $5 \leqslant x \leqslant 15$	1. $n = 20$	1. $\sigma^2 = 4$
2. $\beta_2 = 1.0$	2. $5 \leqslant x \leqslant 85$	2. $n = 40$	2. $\sigma^2 = 9$
3. $\beta_3 = 1.5$			3. $\sigma^2 = 16$
			4. $\sigma^2 = 25$
			5. $\sigma^2 = 36$
			6. $\sigma^2 = 49$
			7. $\sigma^2 = 64$
			8. $\sigma^2 = 81$

(96 possible combinations)

Although the generation of the independent, normally distributed random observations V_i with desired variance would appear conceptually easy, in actuality it involves a considerable amount of justification. The reader interested in this problem and its subtleties is referred to (*13*, pp. 39-61).

For each combination of sample characteristics and the corresponding unique deterministic data set, a total of 100 different sets of independent error values based on the error variance are created and added. The result is 100 different data sets for each of the 96 combinations. Applying least squares analysis to each set yields 100 estimates of β_1 for each of the 96 possible sample combinations enumerated in table 1. (The total is 9,600 regressions, not including one replication for $\sigma^2 = 16$). Each of these sets of 100 estimates constitutes, under slightly different conditions, a small empirical distribution of the estimator $\hat{\beta}_1$, and each provides a sample for estimating $E(\hat{\beta}_1) = \beta_1 +$ bias. Since we know the actual value of β_1, we can subtract it from our estimate of β_1, namely, $\bar{\beta} = 1/100 \sum_{i=1}^{100} \hat{\beta}_{1i}$. The difference, b_i, provides a good estimate of the bias created in estimating β_1 under conditions of error misspecification. The $\bar{\beta}$ are presented in table 2 for each of the respective combinations.

A similar process was followed in generating Monte Carlo results for functional form II, $Y_i + \beta_0 X_{1i}^{\beta_1} X_{2i}^{\beta_2} + V_i$. However, the 1,444 total combinations were based on fewer error variance levels and dual β-values and X-range values (table 3). The $\bar{\beta}$ corresponding to each combination are shown in table 4 (14,400 regressions).

HYPOTHESES OF INTEREST

In examining the results of the Monte Carlo experiment a number of hypotheses are of interest. Those listed below that do not have an alternative hypothesis explicitly stated should be assumed to have, for an alternative hypothesis, H_a: There exists at least one inequality. Bias refers to the estimates of the respective β's.

1. H_0: All b_{ijkl}/s_{ijkl} are equal[2]
2. H_0: The biases due to the levels of "return to scale" averaged over all other factors are equal
3. H_0: The biases due to the different sample sizes averaged over all the other factors are equal
4. H_0: The biases due to the different "regions of the independent variables" averaged over all other factors are equal
5. H_0: The biases due to the different "error variances," averaged over all other factors, are equal
6. H_0: There are interactions in the factors of the experiment

vs H_a: There are no interactions in the factors of the experiment

The hypothesis stated in 6 is ambiguous because a number of different interaction hypotheses can be stated. At this point, it is unnecessary to state them specifically; it is only essential to point out that such hypotheses need special consideration.

If equality is accepted in these hypotheses, it is pertinent to ask if the common value is different from zero and, if so, in what direction is it different. If equality is rejected, then obviously some of the biases are different from zero and only the differences need to be investigated.

It can be shown that rejection of any of the hypotheses, 2, 3, 4, 5, and 6, will not indicate the acceptance or rejection of hypothesis 1. If there is interest in other hypotheses, the model must be reclassified and an additional analysis of variance performed.

A completely randomized design involving a factorial treatment arrangement was taken from statistical theory to assess the general impacts and to test hypotheses 2 through 6. Cost factors in generation of the data have prevented obtaining a complete replicate of the experiment that would be necessary to test hypothesis 1.

---

[2] This is a standardizing transformation of the $\hat{\beta}_{ijkl}$ to satisfy assumptions of ANOVA factorial models. A more detailed justification appears later in the article.

Table 2.—Monte Carlo results for functional form I

			$\beta = .5$		$\beta = 1$		$\beta = 1.5$	
		Variance	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation
$5 \le x \le 15$	n = 20	4	.5160	.1373	.9959	.0508	1.4995	.0191
		9	.5214	.2156	1.0137	.0715	1.5010	.0267
		16	.5393	.2709	.9995	.0914	1.5025	.0376
		16	.5560	.2952	1.0113	.0950	1.5013	.0383
		25	.5429	.3855	1.0337	.1308	1.5060	.0524
		36	.5332	.5973	1.0713	.1689	1.5135	.0575
		49	.5667	.6756	1.0295	.2114	1.5201	.0761
		64	.4736	.5968	1.0195	.2273	1.4980	.0796
		81	.4268	.5701	1.1304	.3394	1.5039	.0869
	n = 40	4	.5134	.0945	1.0075	.0352	1.5001	.0144
		9	.5324	.1554	1.0163	.0590	1.5013	.0207
		16	.5215	.2344	1.0010	.0684	1.5022	.0266
		16	.5854	.2319	1.0196	.0714	1.5061	.0278
		25	.5993	.3348	1.0170	.0904	1.5039	.0319
		36	.5377	.3163	1.0377	.1124	1.5091	.0456
		49	.4646	.3410	1.0582	.1358	1.5162	.0422
		64	.4825	.4285	1.0681	.1744	1.5060	.0561
		81	.3916	.4183	1.0596	.1859	1.5158	.0682
$5 \le x \le 85$	n = 20	4	.5100	.0512	1.0015	.0130	1.4996	.0042
		9	.5191	.0795	1.0035	.0182	1.5011	.0058
		16	.5201	.0953	1.0030	.0285	1.5010	.0075
		16	.5340	.1108	1.0020	.0234	1.4993	.0083
		25	.5334	.1278	1.0013	.0279	1.5007	.0096
		36	.5636	.1810	1.0075	.0451	1.5004	.0123
		49	.5876	.2075	1.0039	.0500	1.5002	.0133
		64	.5411	.1762	1.0126	.0529	1.5040	.0172
		81	.5555	.2275	1.0041	.0610	1.5016	.0184
	n = 40	4	.5029	.0274	1.0010	.0093	1.4999	.0031
		9	.5098	.0526	1.0006	.0147	1.5000	.0054
		16	.5324	.0800	1.0011	.0237	1.4991	.0067
		16	.5151	.0625	1.0039	.0205	1.4995	.0073
		25	.5537	.1111	1.0077	.0250	1.5009	.0088
		36	.5548	.1149	1.0064	.0299	1.5029	.0107
		49	.5556	.1330	1.0029	.0364	1.5032	.0118
		64	.5526	.1415	1.0134	.0412	1.5009	.0137
		81	.5610	.1806	1.0058	.0477	1.4996	.0149

Table 3.—Sample characteristics for functional form II

β values	X range	Sample size	Error variance
1. $\beta_1 = .25$ $\beta_2 = .25$	(5x15) x (5x15)	$n = 20$	$\sigma^2 = 9$
2. $\beta_1 = .4$ $\beta_2 = .1$	(5x15) x (5x85)	$n = 40$	$\sigma^2 = 25$
3. $\beta_1 = .5$ $\beta_2 = .5$	(5x85) x (5x85)		$\sigma^2 = 49$
4. $\beta_1 = .9$ $\beta_2 = .1$			$\sigma^2 = 81$
5. $\beta_1 = .75$ $\beta_2 = .75$			
6. $\beta_1 = 1.4$ $\beta_2 = .1$			

(1,444 possible combinations)

EMPIRICAL RESULTS

Empirical Distributions and ANOVA Models

Consider the data layout in tables 2 and 4, in which the entries are $\bar{\beta}$ values. Each is in a cell characterized by four subscripts, i, j, k, and l, where i denotes the i^{th} "return to scale," j denotes the j^{th} "region of the independent variable," k denotes the k^{th} "sample size," and l denotes the l^{th} "error variance."

From an examination of the data layout, it is apparent that a completely randomized design involving a factorial treatment arrangement should be used. For functional form I, ANOVA will be applied, and, recognizing the presence of two responses under form II, MANOVA techniques will be used. For more information on both techniques, see (*7, 8, 10*). The distribution of $\bar{\beta}_{ijkl}$ indicates that certain transformations must be made to these data to realistically satisfy the normality of the factorial model. Since $\bar{\beta}_{ijkl}$ is a mean of 100 estimates, $\bar{\beta}_{ijkl}$ is normally distributed with variance $\frac{\sigma^2_{ijkl}}{100}$, where σ^2_{ijkl} is the variance of $\hat{\beta}_{ijkl}$. The variance of $\hat{\beta}_{ijkl}$ is not known, but a good estimate of $\sigma^2 ijkl$ based on 100 observations is s^2_{ijkl}. So $\bar{\beta}_{ijkl}$ is distributed approximately $N(\beta_i + b_{ijkl}, \sigma^2_{ijkl}/100)$ and, consequently, $(\beta_{ijkl} - \beta_i)/s_{ijkl}$ is approximately distributed $N(b_{ijkl}/s_{ijkl}, 1/100)$.

Since there are no *a priori* reasons for making strong assumptions concerning interaction and there is generally only one observation per cell, the analysis will proceed in stages. The first stage will be to use a three-way factorial analysis of variance model on a transformation of the data in table 2 that corresponds to:

- The three levels of "return to scale" $\beta = .5$, 1.0, and 1.5;
- The two "regions of the independent variable" $5 \leqslant X \leqslant 15$, $5 \leqslant X \leqslant 85$;
- The two "sample sizes" $n = 20$, $n = 40$.

The "error variance" is $\sigma^2 = 16$ with two replications.

Based on the conclusions of the first stage, we will use a modified four-way factorial model on the transformed data excluding only that part of the data that was generated to use as a replicate in the factorial model of stage one.

On the basis of the conclusions of the first two stages, the selection of a multivariate factorial model for analyzing the second functional form $Y_i = \beta_0 X_{1i}^{\beta_1} X_{2i}^{\beta_2} + V_i$ can be made. The results of the Monte Carlo application to this functional form appear in table 4.

Functional Form I $Y = \beta_0 X^{\beta_1}$

In the first stage the concern is the degree of interaction of effects among the sample characteristics considered:

1. Return to scale $= \alpha_i$ $(i = 1, \ldots, 3)$
2. Region of independent variable $= \nu_j$ $(j = 1,2)$
3. Sample size $= \delta_k$ $(k = 1,2)$
4. Error variance $= \delta_l$ $(l = 1, \ldots, 9)$

To find out if there are interactions among the characteristics that will require making allowances for them, the analysis begins by holding the error variance fixed and varying factor 1, factor 2, and factor 3 where the assumed model is

$$Y_{ijkm} = \mu + \alpha_i + \upsilon_j + \gamma_k + (\alpha\upsilon)_{ij} + (\alpha\gamma)_{ik} + (\upsilon\gamma)_{jk} + (\alpha\upsilon\gamma)_{ijk} + \epsilon_{ijkm} \qquad (12)$$

$i = 1, 2, 3 \qquad j = 1, 2 \qquad k = 1, 2 \qquad m = 1, 2$

Since this model provides no estimate of error if there is no replication, two sets of $\bar{\beta}$ values were obtained for the error variance $\sigma^2 = 16$. Consequently, we have a three-way factorial model with two observations per cell.

It is not realistic to assume that the ϵ's are distributed independently $N(0, \sigma^2)$. More realistically, we should assume that $\underline{\epsilon} \sim N(\underline{0}, \sigma^2\Sigma)$, where Σ is a diagnonal matrix. Consequently, the $\bar{\beta}$ values which appear in the tables have been transformed for ANOVA by subtracting out the actual value of the parameter and dividing by the estimated standard deviation.

Table 4.—Monte Carlo results for functional form II

		σ_ν^2	$\beta = .25$		$\beta = .25$		$\beta = .4$		$\beta = .1$		$\beta = .5$		$c = .5$	
			Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation
5x15 x 5x15	n = 20	9	.3107	.2050	.2889	.2176	.4151	.1674	.1061	.1947	.4950	.0669	.4984	.0568
		25	.3281	.4320	.2391	.4595	.4504	.3980	.1874	.3705	.5131	.1213	.5019	.1002
		49	.2296	.5820	.3609	.5020	.4314	.5992	.1893	.5359	.5056	.1686	.5337	.1387
		81	.2367	.6684	.2203	.5896	.2713	.5980	.0635	.6270	.5364	.2364	.5389	.2217
	n = 40	9	.2811	.1500	.2770	.1473	.4289	.1621	.1297	.1436	.5050	.0468	.5054	.0479
		25	.2544	.2738	.2920	.2972	.3973	.2811	.2031	.2488	.5161	.0811	.5069	.0714
		49	.2783	.3626	.2225	.3834	.2918	.3817	.1512	.4389	.5383	.1299	.5142	.1263
		81	.2815	.4328	.2100	.4513	.3086	.4532	.1609	.4001	.5429	.1563	.5446	.1472
5x15 x 5x15	n = 20	9	.2483	.0708	.2534	.1292	.4168	.0813	.0937	.1109	.5027	.0204	.4989	.0428
		25	.2708	.1588	.2591	.2411	.4275	.1318	.1153	.2119	.5035	.0325	.5033	.0638
		49	.2846	.2163	.2894	.3713	.4138	.1871	.1034	.3632	.5052	.0491	.4948	.0910
		81	.3085	.2960	.3020	.4954	.4508	.2279	.1133	.4389	.5136	.0623	.5146	.1067
	n = 40	9	.2589	.0458	.2764	.1138	.4102	.0514	.1045	.0888	.5020	.0129	.5014	.0249
		25	.2702	.1061	.2702	.2025	.4387	.1175	.1025	.1642	.5012	.0233	.5032	.0454
		49	.2748	.1356	.3041	.2750	.4604	.1557	.0740	.2388	.5077	.0378	.5043	.0570
		81	.2618	.1853	.2786	.3963	.4372	.1629	.1022	.2997	.5163	.0450	.5247	.0839
	n = 20	9	.2556	.0560	.2594	.0545	.4099	.0693	.0988	.0603	.4999	.0125	.5014	.0135
		25	.2734	.1017	.2523	.1013	.4256	.1217	.1077	.1134	.5052	.0220	.5007	.0198
		49	.2859	.1721	.2948	.1559	.4658	.1883	.1100	.1427	.5096	.0366	.5052	.0328
		81	.2943	.2146	.2755	.1951	.4207	.1790	.0897	.1965	.5067	.0425	.4985	.0402
	n = 40	9	.2538	.0430	.2549	.0437	.4071	.0449	.1019	.0371	.4997	.0060	.4943	.0096
		25	.2604	.0707	.2591	.0709	.4254	.1157	.1135	.0791	.5003	.0119	.5009	.0158
		49	.2803	.1047	.2955	.1305	.4438	.1361	.0858	.1177	.5031	.0196	.5043	.0238
		81	.2679	.1557	.2734	.1547	.4600	.1469	.1061	.1364	.5030	.0254	.5120	.0332

		σ_ν^2	$\beta = .9$		$\beta = .1$		$\beta = .75$		$\beta = .75$		$\beta = 1.4$		$\beta = .1$	
			Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation	Mean of βs	Standard deviation
5x15 x 5x15	n = 20	9	.9129	.0765	.1090	.0533	.7509	.0230	.7529	.0216	1.4064	.0266	.1022	.0204
		25	.9113	.1176	.1016	.1053	.7574	.0383	.7535	.0324	1.4082	.0522	.0952	.0383
		49	.9552	.2524	.1182	.1621	.7494	.0533	.7506	.0463	1.4100	.0640	.0926	.0494
		81	.9859	.2462	.1249	.1836	.7643	.0755	.7589	.0607	1.4050	.0794	.0938	.0605
	n = 40	9	.9122	.0569	.0963	.0410	.7477	.0151	.7511	.0159	1.3981	.0220	.0983	.0165
		25	.9214	.0844	.1026	.0654	.7516	.0290	.7538	.0328	1.4066	.0344	.0963	.0246
		49	.9171	.1343	.1110	.0971	.7560	.0329	.7593	.0387	1.4041	.0482	.1022	.0362
		81	.9606	.1964	.1492	.1941	.7518	.0569	.7566	.0539	1.4066	.0676	.1094	.0418
5x15 x 5x85	n = 20	9	.9033	.0219	.0953	.0294	.7500	.0063	.7511	.0103	1.4000	.0060	.0997	.0073
		25	.9009	.0312	.0963	.0419	.7488	.0112	.7494	.0174	1.4002	.0111	.1002	.0138
		49	.9095	.0462	.0897	.0690	.7509	.0159	.7518	.0312	1.4013	.0127	.0997	.0158
		81	.9142	.0733	.0849	.0955	.7525	.0190	.7501	.0305	1.4029	.0208	1.0161	.0269
	n = 40	9	.9014	.0143	.0990	.0158	.7496	.0041	.7505	.0076	1.4004	.0049	.0991	.0038
		25	.9005	.0229	.1053	.0262	.7503	.0079	.7504	.0143	1.4001	.0095	.0990	.0064
		49	.9095	.0340	.0974	.0372	.7497	.0087	.7508	.0160	1.4030	.0107	.1003	.0082
		81	.9144	.0585	.1018	.0507	.7524	.0129	.7505	.0251	1.4015	.0157	.0996	.0138
5x85 x 5x85	n = 20	9	.9006	.0239	.1007	.0112	.7504	.0032	.7500	.0031	1.3984	.0096	.0995	.0031
		25	.9061	.0477	.0985	.0218	.7501	.0052	.7503	.0053	1.3998	.0173	.0993	.0049
		49	.9099	.0618	.1014	.0306	.7504	.0091	.7502	.0065	1.4059	.0223	.0996	.0070
		81	.9233	.0796	.1062	.0354	.7508	.0096	.7495	.0087	1.4017	.0306	.1001	.0081
	n = 40	9	.9008	.0146	.1000	.0087	.7496	.0026	.7497	.0018	1.3995	.0049	.1000	.0028
		25	.8999	.0281	.0995	.0171	.7501	.0043	.7500	.0031	1.4002	.0077	.0997	.0044
		49	.9035	.0313	.1036	.0212	.7500	.0066	.7502	.0049	1.4010	.0102	.0998	.0054
		81	.9167	.0466	.0972	.0300	.7510	.0084	.7506	.0051	1.4001	.0152	.0994	.0078

The analysis of variance results in table 5 indicate that no significant interaction effects of the second or third order occur in the three variables "return to scale," "sample size," and "region of independent variable." There exists the possibility that a "return to scale - region of independent variable" interaction exists, since the F-ratio of 2.28 would be significant at a probability level of approximately .175. The F-ratio 7.88 indicates a probability greater than .99 that "return to scale" affects the bias.

Table 5.—ANOVA for a three-factor factorial with replication: functional form I

Factor	Degrees of freedom	Sum of squares	Mean squares	F-ratio
1	2	.18863	.09431	[a]7.88
2	1	.00124	.00124	.10
3	1	.00716	.00916	.60
12	2	.05454	.02727	2.28
13	2	.00676	.00338	.28
23	1	.01348	.01348	1.13
123	2	.01400	.00700	.58
Error	12	.14366	.01197	
Total	23	.42947		

[a]Significant at .01 level.

The means for the levels of "return to scale" are .2449, .1051, and .0310 for $\beta = .5$, $\beta = 1$, and $\beta = 1.5$, respectively. Hence, the bias is greatest when the return to scale is lowest and least when the return to scale is greatest. The means for the levels of "return to scale" are independent, unbiased estimates of the biases due to their respective levels. In theory, the error variance in the ANOVA is independent of the means; consequently, Tukey's T method can be used as a followup analysis to determine the nature of the differences in levels of bias corresponding to the levels of "return to scale." The T method indicates that the mean response for $\beta = .5$ and for $\beta = 1$ differ at approximately the .06 significance level and the mean response for $\beta = .5$ and for $\beta = 1.5$ differ at the .01 significance level.

Since the "return to scale" factor is quantitative, we can consider the functional form of the response curve for this factor. When there are just three treatments, the sum of squares can only be split into two quantities, linear and quadratic responses. The total sum of squares for "return to scale" is .18863, and the sum of squares for linearity and for a quadratic response are .18286 and .00577, respectively. This would suggest a linear response curve.

Since the results of the ANOVA suggest that a significant interaction exists between "return to scale" and "region of independent variable" we now incorporate it and a fourth variable "error variance" into the model. The new model is

$$Y_{ijkl} = \mu + \alpha_i + \upsilon_j + \gamma_k + \delta_l + (\alpha\upsilon)_{ij} + (\alpha\delta)_{il} + (\upsilon\delta)_{jl} + (\gamma\delta)_{kl} + \epsilon_{ijkl} \qquad (13)$$

$i = 1, 2, 3 \qquad j = 1, 2 \qquad k = 1, 2 \qquad l = 1, 2, \ldots, 8$

Note that third and fourth order interactions do not appear in the model. There are two reasons: (1) in the first stage, higher order interactions seemingly did not exist, and (2) each higher order interaction that can be safely assumed out of the model causes an increase in the stability of the estimate of the error variance. It should also be noted that the significant interaction effect found in the first stage, $(\alpha\upsilon)_{ij}$, remains in this model.

The distribution of $\underline{\epsilon}$ is again most realistically assumed to be $N(\underline{0}, \sigma^2\Sigma)$; consequently, the $\overline{\beta}$ must be transformed for the analysis.

The results of the analysis of variance for the second stage are presented in table 6.

Table 6.—ANOVA for a four-factor factorial without replication: functional form I

Factor	Degrees of freedom	Sum of squares	Mean squares	F-ratio
1	2	.15087	.07544	[a]5.825
2	1	.04364	.04364	3.370
3	1	.04205	.04205	3.247
4	7	.29436	.04205	[a]3.247
12	2	.53364	.26682	[a]20.604
14	14	.24608	.01758	1.357
24	7	.06162	.00880	.680
34	7	.05405	.00772	.596
Error	54	.69913	.01295	
Total	95	2.12544		

[a]Significant at .01 level.

Table 6 F-ratios indicate significant main effects for "return to scale" and "error variance" and a very significant "return to scale - region of independent variable" interaction.

The presence of the interaction complicates the making of inferences concerning main effects. Mean values for a "return to scale" across "region of independent variable" classification appear below:

Data range	$\beta = .5$	$\beta = 1.0$	$\beta = 1.5$
$5 < x < 15$	.0568	.2233	.1192
$5 < x < 85$	.3090	.1408	.0775

The value that conspicuously contributes to the interaction effect is .0568. This mean is the result of averaging 16 $\bar{\beta}$ values generated when $\beta = .5$, $5 < x < 15$, $n = 20$, $n = 40$ and $\sigma = 2, 3, \ldots, 9$. When σ took on the values 7, 8, and 9, the distribution of the error term at the lower end of the region $5 < x < 15$ was a truncated normal distribution. The truncation was in the neighborhood of 1 standard deviation away from the mean. To illustrate, consider the error distribution when $x = 5.0625$ and $\sigma = 9$. Since $Y = 4X^{.5} = 9.00$, it is clear that if negative values were allowed, the distribution of $Y = 4X^{.5} + \epsilon$ would be normal with mean 9 and variance 81. However, the logarithm transformation demands positive Y values, so the generation process would not allow any error values smaller than -9. Thus, the error density at $\sigma = 9$ is

$$f(\epsilon) = \left\{ [1-F(-9)]\, 3\sqrt{2\pi} \right\}^{-1} \exp\left\{ -\epsilon^2/18 \right\}.$$

Such an error distribution will tend to increase the level of the Y values in the lower part of the region of the independent value. The problem is not as severe at the upper end of the region since the deterministic portion of Y is larger. Consequently, the estimated slope in the linear equation is lower than it would be without the truncation problem.

The truncation difficulties can be partially removed by limiting the error standard deviation to values less than 6. Under this subset the mean values become:

Data range	$\beta = .5$	$\beta = 1$	$\beta = 1.5$
$5 < x < 15$	.1501	.1332	.0584
$5 < x < 85$	.2610	.1203	.0296

Even with this restriction of the data, the interaction effect is still present. Thus, the general inferences remain valid, namely that analysis of the main effects due to "return to scale" must be tempered by the effects of the region of the independent variable.

The overall means for $\beta = .5$, $\beta = 1$, $\beta = 1.5$ are respectively .1829, .1820, and .0984. The followup comparison, again using Tukey's T method, indicates the biases for $\beta = .5$ and $\beta = 1$ are not distinguishable, but both are different from the bias for $\beta = 1.5$ at the .01 level. The decrease in error for larger "return to scale" values, particularly when the X region is larger, results from two factors. First, as the error variance decreases, the model becomes deterministic and the logarithmic transformation becomes more appropriate. Second, when the deterministic portion of Y dominates the standard error, there is a negative bias (as pointed out earlier). This bias should tend to counteract the positive bias reflected in the earlier stages of this analysis.

The means for the error variances 4, 9, . . . , 81 are, respectively, .0640, .1438, .0977, .1971, .2399, .2080, .1617, and .1232. An orthogonal breakdown for the "error variance" sum of squares appears below:

linear	.05500
quadratic.	.15779
cubic	.01507
remainder	.06650
total	.29436

This indication of a quadratic response for "error variance" reflects the truncation difficulty mentioned earlier. As the error variance increases, the misspecification of the error term becomes increasingly serious. When the error standard deviation effectively overshadows the deterministic portion of Y that corresponds to small values of X, the general level of the Y observation increases and pushes the estimates of β in the opposite direction. Hence, the bias goes down.

The T method followup indicates that adjacent values are not distinguishable at the .1 level.

Functional Form II $\quad Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2}$

Analyzing the empirical distributions in table 4 for functional form II presents a new difficulty. Functional form II, $Y_i = \beta_0 X_{1i}^{\beta_1} X_{2i}^{\beta_2} + V_i$, has two "return to scale" parameters, and as a consequence, the data are in two dimensional vector form. It cannot be assumed that the estimates of β_1 and β_2 will be independent; therefore, it will be necessary to use multivariate analysis of variance (MANOVA) techniques to analyze the consequence of misspecifying the error term.[3]

The results, using a four-way multivariate factorial analysis of variance model with two responses, are presented in table 7.

From the MANOVA table the following observations can be made:

- "Return to scale" main effect significant at the .01 level
- "Range of the X variables" main effect significant at the .01 level
- "Error variance" main effect at the .10 level
- "Return to scale-range of the X variable" interaction significant at the .01 level
- "Return to scale-error variance" interaction significant at the .10 level
- "Range of the X variable-error variance" interaction significant at the .10 level

The presence of the interactions, of course, complicates a straightforward assessment of the main effects. It does seem reasonable, however, that the "return to scale"

---

[3] There are several possible tests for analyzing data under MANOVA assumptions. The likelihood ratio tests are used in this study. For a complete treatment see (*1*, chapter 8).

Table 7.—MANOVA for a four-factor factorial: functional form II

Factor	U-statistic	Approximate F-statistic	Degrees of freedom	
1	0.250339	^{a}5.7921	10	58
2	.629761	^{a}3.7717	4	58
3	.912625	1.3882	2	29
4	.687525	^{b}1.9916	6	58
12	.226585	^{a}3.1923	20	58
13	.911772	.2741	10	58
23	.906933	.7258	4	58
14	.305862	^{b}1.5624	30	58
24	.539751	^{b}1.7455	12	58
34	.888048	.5912	6	58
123	.499383	1.2038	20	58
124	.338052	.6959	60	58
134	.524836	.7353	30	58
234	.872698	.3405	12	58

^{a}Significant at the .01 level. ^{b}Significant at the .10 level.

factor will affect the bias and, further, that the "region of the X variables" would similarly affect the bias in the estimates of return to scale. It is not clear, however, that changing the error variance will affect the bias as long as the ratio of the deterministic part of Y and the error standard deviation remains at a reasonable level.

The error standard deviation may be so large that a truncated error distribution is effectively at work on the lower end of the deterministic curve. If so, certainly the "error variance" will affect the bias of the estimates of return to scale. The mean vectors for error variances 9, 25, 49, and 81, respectively, are [.0788, .0277], [.1068, .0397], [.1437, .0673], and [.1507, .0745]. Realizing the effect of the truncated error distribution and consequently allowing for its effect, we can note that the means for 49 and 81 error variances should be somewhat greater than they are. Hence, discounting the truncation factor, the "error variance" main effect will be more significant, and increasing the error variance does increase the bias resulting from error misspecification.

Table 8 gives the mean vectors associated with the "return to scale" and "region of the X variables" classification. To see the interaction effect in table 8, note the behavior of the vectors in the two "wide range" rows as the "return to scale" increases. Contrast this behavior with the behavior of the vectors in the "narrow range" row as the "return to scale" increases. This interaction effect can be partially explained by the truncation problem and partially by the domination of the standard deviation of the error by the deterministic part of Y for the higher "return to scale" values. Hence, this interaction effect is at least somewhat inherent in the misspecification of the error.

The cell means for the "range of the X variable" levels 5x15 x 5x15, 5x15 x 5x85, and 5x85 x 5x85, are respectively, [.1062, .0931], [.1202, .0436] and [.1337, .0315]. The vector [.1062, .0931] is distinguishable from the other two vectors. And although some interaction effects are immersed in these means, it seems that areas of limited size for the "range of the X variables" will increase the bias in estimating the return to scale when the error is misspecified. Note that to speak of the size of the region is somewhat misleading, since the principal determinant of the effect is the level of the deterministic portion of the curve.

The cell means for the subsets of "return to scale" levels .5, 1.0, and 1.5 are, respectively, [.1365, .1175], [.1467, .0668], [.1388, .1009], [.1701, .0233], and [.0527, .0529], [.0751, -.0475]. An examination of

Table 8.—Mean vectors associated with return to scale and region interaction

Region of independent variables	Returns to scale					
	$\beta_1 = .25$ $\beta_2 = .25$	$\beta_1 = .4$ $\beta_2 = .1$	$\beta_1 = .5$ $\beta_2 = .5$	$\beta_1 = .9$ $\beta_2 = .1$	$\beta_1 = .75$ $\beta_2 = .75$	$\beta_1 = 1.4$ $\beta_2 = .1$
5x15 x 5x15	.0995 .0738	-.0328 .1728	.1369 .1293	.2164 .0935	.1040 .1187	.1131 †.0297
5x15 x 5x85	.1370 .1095	.2322 .0024	.1671 .0753	.1530 -.0564	.0216 .0371	.0910 -.0465
5x85 x 5x85	.1731 .1692	.2408 .0252	.1125 .0983	.1409 .0327	.0325 .0027	.0213 -.0664

these means reveals the overall decrease in bias resulting for higher levels of "return to scale." Thus, seriousness of the bias created by error misspecification for an increasing return to scale is less than for a constant or a decreasing return to scale.

CONCLUSIONS

The factors considered in the foregoing analysis were "return to scale," "region of the independent variables," "sample size," and "error variance." The sample sizes examined gave no indication of having distinguishable effects on estimate bias. The "error variance" does affect the bias and, generally speaking, it causes an increase in the bias as the error variance increases. The "region of the independent variable" affects the bias in the estimates in a couple of ways. First, if the region is small and the level of the deterministic portion of Y does not totally dominate the error standard deviation, then the bias is positive and can be substantial. Second, if the region is large and, consequently, the deterministic portion of Y dominates the error standard deviation for a good portion of the region, the bias effect will decrease. In cases where the deterministic portion dominates almost entirely, the bias becomes quite small and it can be negative. The "return to scale" affects the bias in a manner similar to the "region of the independent variable" in that return to scale affects the deterministic portion of Y. Consequently, as the "return to scale" level increases, it causes the deterministic portion to dominate the error standard deviation, and the bias is not serious. As the "return to scale" level decreases, the bias can be substantial.

These conclusions are by no means unexpected, but one is impressed by the robustness of the least squares technique under error misspecification of this type. A cursory review of tables 2 and 4 indicates that, except under the most difficult conditions with regard to variance, the bias generally is not significant enough to affect the empirical application of the corresponding elasticities. The evidence suggests bias would be present but not of the magnitude, in most cases, to greatly challenge previous economic evaluations based on regressions that use Cobb-Douglas type functions. While this is a source of consolation for many analysts, it should be noted that any conclusion that loglinear estimation yields satisfactory estimates in the presence of additive errors must still depend on the peculiarities of the sampling experiment (*9*, p. 445).

REFERENCES

(1) Anderson, T. W. *An Introduction to Multivariate Statistical Analysis.* John Wiley & Sons, Inc., 1958.
(2) Edwards, Clark. "Non-Linear Programming and Non-Linear Regression Procedures." *Jour. Farm. Econ.*, XLIV, No. 1, February 1962.
(3) Goldberger, A. S. *Econometric Theory.* John Wiley & Sons, Inc., 1963.
(4) Goldberger, A. S. "The Interpretation and Estimation of Cobb-Douglas Functions." *Econometrica*, XXXV, No. 3-4, July-October, 1968.
(5) Goldfeld, Stephen M. and Richard E. Quandt. "The Estimation of Cobb-Douglas Type Functions With Multiplicative and Additive Errors." *International Econ. Rev.*, XI, No. 2, 1970.
(6) Heien, Dale S. "A Note on Log-linear Regression." *Jour. Amer. Statis. Assoc.* Vol. 63, No. 323, September 1968.
(7) Morrison, Donald F. *Multivariate Statistical Methods.* McGraw Hill Book Co., 1967.
(8) Ostle, Bernard. *Statistics in Research.* Iowa State Univ. Press, 1963.
(9) Quandt, Richard E. "The Estimation of Constant Elasticities: Comment." *Southern Econ. Jour.* Vol. XXXIX, No. 3, January 1973.
(10) Scheffe, Henry. *The Analysis of Variance.* John Wiley & Sons, Inc., 1959.
(11) Theil, H. *Principles of Econometrics.* John Wiley & Sons, Inc., 1971.
(12) Theil, H. "Specification Errors and the Estimation of Economic Relationships." *Rev. International Statis. Inst.*, XXV, 1957.
(13) Tocher, K. D. "The Application of Automatic Computers to Sampling Experiments." *Jour. Royal Statis. Soc.*, Series B, Vol. 16, 1954, pp.39-61.
(14) Weber, J. E. and Hawkins, C. A. "The Estimation of Constant Elasticities." *Southern Econ. Jour.*, October 1971, pp. 185-192.

RESEARCH REVIEW

IN THIS ISSUE

"The gods did not reveal to men all things from the beginning," according to a fragment written some 2,600 years ago by the Greek philosopher Xenophenes, "but men through their own search find in the course of time that which is better." To the economist, social progress discussed by Xenophenes is captured statistically in measures of productivity; finding new ways of doing things leads to more output valued by persons per unit of input.

For U.S. agriculture, continued progress has been reflected in the increasing number of persons supplied with food and fiber per farmworker. There were 56 persons supplied per farmworker in 1974, compared with 15.5 in 1950 and 4.1 in 1820. The trouble is that this simple measure completely overlooks the nonfarm inputs that have become increasingly important in our food and fiber system since 1820. An alternative measure of productivity for the food and fiber sector is clearly needed.

In this issue, we offer two such measures. They indicate that about 16 or 17 persons are currently supplied per worker in the total food and fiber sector. One measure relies on the conceptual framework of input-output for which data are available at 5-year intervals with some lag. The other is based on a simpler concept, but data can be provided annually with a short lag. Both measures have advantages and disadvantages. Readers of this journal may want to react to these proposals or suggest further alternatives to the National Economic Analysis Division of ERS.

Empirical results of research in which nonlinear regression analysis is used vary according to assumptions made in specifying the error term. Some researchers give this little or no thought as they choose between nonlinear regression and a linear fit to the logarithms of the variates. A Monte Carlo simulation of the alternatives for fitting a Cobb-Douglas production function, reported in this issue, indicates that the choice may have sharp implications for the ultimate research findings.

Clark Edwards

THE FEAR OF KNOWLEDGE

Intellectuals have generally been held in relatively high esteem in our country, and we have placed heavy emphasis on higher education. The extent of this support, however, has varied sharply over the past 200 years.

Some of the greatest minds of 17th-century America—including Thomas Jefferson, Alexander Hamilton, and James Madison—once collaborated in preparing a speech for George Washington that, although never delivered, said:

> Promote then, as an object of primary importance, institutions for the general diffusion of knowledge. In proportion as the structure of a government gives force to public opinion, it is essential that public opinion be enlightened.

It is with considerable irony that we note that a society founded and nurtured by intellectuals, a society which has benefited so greatly from a spectacular development and utilization of knowledge, also is a society in which a spirit of anti-intellectualism has periodically slowed economic and social development. The early leaders of our country emphasized education and social enlightenment. They reconciled wide-ranging intellectualism with practical political activity. The fear, in that day, was of a lack of knowledge, not of its acquisition and use.

The initial intellectualism in American government lasted until the Jacksonian era, when it became resented as a boorish kind of excellence, as an unjustified claim to distinction, as a challenge to egalitarianism, and as a quality that somehow deprived a man or woman of what was known as the "common touch."

In the years between the Jackson and Lincoln administrations, there was a general intellectual reawakening. Part of this was attributed to the call for new knowledge to enhance industrial and agricultural expansion. There also was a spirit of intellectual liberality exemplified by such men as Emerson, Longfellow, and Hawthorne.

The lyceum movement provided a stimulus to intellectual curiosity, and the land-grant system emerged. The Land-Grant College Act of 1862 was passed during the administration of Abraham Lincoln, a president who exemplified the concept that intellectual development was democratic and egalitarian, and that it should be the universal aspiration of all men and women.

From Lincoln's time until after the turn of the century, our country underwent a great economic and industrial expansion and transformation. Politics and government were oriented toward the acquisition of wealth and power. Reorientation and growth occurred in our educational institutions. Indeed, the first thrust toward reform in scandal-ridden government in the late 1800's began on the college campuses, just as it did in the late 1960's.

Theodore Roosevelt was an intellectual who recognized that, in his time, only a leader who appealed to the people on the basis of physical and mental characteristics could be successful. Roosevelt was exceedingly well read and well informed. He initiated reforms that strengthened the Nation's economic development by changing its orientation toward the common good, rather than toward the benefit of the privileged few.

Following Woodrow Wilson's administration, the country went through the intellectual drought of the twenties and thirties. The Ku Klux Klan gained strength. The Scopes Trial comprised an all-out attack on intellectualism. And, the Depression years discouraged higher education.

Then came the administration of Franklin Roosevelt, and the country, for the first time since the days of Jefferson and Hamilton, saw a major influx of intellectuals into Federal service. They aroused much antagonixm but they changed the shape and direction of Government.

The pendulum again swung full circle when, in the early part of the fifties, the country was subjected to the "great inquisition" of Joseph McCarthy. Probably the most intense surge of anti-intellectualism since the Jacksonian era, it subjected intellectuals in Government to a frustrating suspicion.

The "GI" Bill had a major impact upon higher education following World War II, making opportunities for higher education a reality for millions of young men and women. But the educational system in America did not progress as rapidly as it did in some other countries. We became painfully aware of our complacency about education when the implications of Sputnik became obvious. Immediately, Dwight Eisenhower moved to improve our educational system.

John Kennedy gave strength and purpose to intellectual ideals. The stream of academics to Washington during this period was great. The Kennedy administration expanded the Nation's educational programs and committed vast amounts of new funding to scientific and technological research. Faced with the commitment to solve complex problems and the necessity of having access to the best available knowledge, the Government turned increasingly to the universities. Among academicians there was a growing feeling that knowledge was useful, that new knowledge was needed, and that the search for it would be supported.

Yet, the American people are impatient. Expectations were developed that could not be fulfilled. Many of the goals proved elusive. Enchantment with research diminished and investment in research was increasingly questioned. While the results of research formed the guiding assumptions for many ventures by Government, the ties between the ideas of the scholars and public policy often proved to be tenuous. There was a distinct difference in the orientation of the scientist and that of the policymaker. The strain between academicians and policymakers was magnified by a serious mismatch between the time required to carry out creative research and the attention span of politicians. As the decade of the 1960's drew to a close, the ties between the intellectuals and the Government had weakened substantially. During Richard Nixon's administration, the President's Science Advisory Council was abolished. Increasingly, people questioned the utility of knowledge.

Today, attacks against knowledge are coming from the left as well as from the right. The recurring resentment against intellectuals, which we have come to expect in our history, has been transformed into a fear of new knowledge. It is sometimes expressed in a wistful desire to return to simpler days. Intellectuals are resented for the changes they have created and feared for those they will create in an uncertain future. More knowledge is equated with the decline in moral standards, the disasters of runaway technology, and the growing complexities of all aspects of human life.

Perhaps even more significant is the growing concern of scientists and academicians over the expansion of knowledge. An increasing number are questioning whether we wish to know more. The fear of knowledge, however, is misplaced. Our fear is not of knowledge, *per se*. It is a fear of how knowledge would be used. Underlying this fear, therefore, is a concern for our own moral weakness and for the possibility that knowledge will not be used for the benefit of society.

What are the lessons to be gleaned from the changing relationships between intellectuals and the public? First, what our society expects of its universities remains unclear. The utilitarian ideas of Franklin and Jefferson concerning the university's services to the community and to the Nation have not been universally accepted either on university campuses or in government. There is, however, a general feeling that the interrelated functions of research, instruction, and public service should continue. Increasing emphasis is being placed upon innovative ways of testing ideas and making the results available.

Second, the knowledge base of the university is lacking relative to the complexity of the problems faced. Our academic base has been noticeably detached from the world around us, giving rise to the outcry of students and the public for more relevancy in higher education. Most professors contend that teaching students how to think is infinitely more important than teaching them what to think about. Emphasis upon theory continues to be necessary in teaching how to acquire knowledge, hypothesize about it, assess it, and integrate it into our thinking. But it is increasingly apparent that students and the public generally desire a better balance between theoretical and empirical analysis.

Third, we are becoming more mindful of unanticipated consequences of new knowledge developed under controlled conditions but applied under general conditions. Concern over research and technological development derives in part from a growing feeling that important secondary effects often are not given due consideration in policy decisions.

Fourth, public policy is controversial. Its essence involves conflict, compromise, and consent. The free flow of new knowledge should help to create citizen awareness of issues and participation in democratic processes. But, given the same data, analyses, and scientific conclusions, individuals and groups may arrive at different decisions.

Fifth, university programs may be out of phase with public support and public needs at any time. The scientist must reserve the right to fail. The policymaker, on the other hand, is not likely to be accorded this privilege. It is important, therefore, that the role of the academician serving as policy advisor be clearly differentiated from that of the policymaker, and that the public continue to support the universities during periods of estrangement between scientists and government.

Sixth, the academician must link the university to other institutions. A gulf has existed between the academician and those who make public policy.

The problems of the complex society in which we live require that the knowledge base be expanded and that access to the best knowledge available be facilitated. The trend toward alienation of the academic community from the policy arena must be halted and reversed. The initiative for this reversal must come from those who are committed to the life of the mind. Only in this way can the stream of knowledge coming from the universities be perpetuated and integrated into the central institutional system of our country.

This note summarizes the ERS Bicentennial Lecture delivered by Charles E. Bishop on June 15, 1976. Dr. Bishop is president of the University of Arkansas.

POLICIES FOR EQUITABLE GROWTH IN DEVELOPING COUNTRIES

After two decades of attempts at raising per capita GNP in developing countries, the development community has recently shifted its focus to the challenge of increasing the equity of income distribution. This shift in emphasis was needed. Empirical studies showed that benefits from economic growth did not trickle down. More serious, economic growth has sometimes resulted in a systematic worsening (both relative and absolute) in the position of persons receiving the least amount of income.

The accompanying table tests major growth and equity hypotheses generated by each of three recent studies:

1. A cross-sectional, statistical analysis (AM in the table) of the sources of differences among countries in the relative amount of income received by the poorest 60 percent of households[1]
2. A historical analysis (Hist in the table) of processes and initial conditions leading to extreme poverty in 24 countries in the middle of the 19th century[2]
3. A model (A-R in the table) of the South Korean economy developed and used to explore the effects[3] of various major (but nonrevolutionary) strategies, policies, and programs on poverty.

In the following table, the interaction with income distribution of each variable or influence is described; the studies on which the description is based are noted; and the reasons for the impact on income distribution are summarized.

All three studies differ completely in methodology and in the settings in which the relationships between poverty and growth are investigated. Yet, all three lead to the following remarkably consistent and reinforcing set of policy conclusions:

Hypotheses concerning the impact of economic processes on income distribution

Variable or influence	Impact and supporting studies	Reasons for impact
Per capita GNP and level of socioeconomic development	For poorest 60 percent of population, their relative share shows U-shaped relationship to per capita GNP (A-M). Relationship between absolute per capita income and income distribution complex; very large differences in development levels positively related to both average income and reduction in poverty; but rapid growth tends to increase poverty even when average income rises (Hist and A-M).	At very low levels of development, a small number of growth points leads to concentration of benefits of economic change in hands of oligarchy of merchants, industrialists, and plantation owners; indirect effects of economic change tend to hurt the very poor by displacing and marginalizing their activities. Only beyond a threshold determined by extent and spread of expansion of economic opportunities is the trickle-down effect sufficient to raise average incomes of urban wage earners and agricultural laborers; even then, extreme poverty is reduced very slowly because the labor market is segmented; human skills lack adaptability.
Short-term rate of growth of per capita GNP	No simple association with income distribution; structure and composition of the increase in income, not average degree of income change, determines impact (A-M and A-R).	Nature, extent, and rapidity of *structural* change govern the direction and magnitude of the net balance of the processes of displacement, absorption, and social adaptation.
Economic innovations	At low and medium-sized levels of development, net impact of structural change of any kind on the poor is systematically unfavorable, even when average incomes rise.	The poor do not have resources and skills to take advantage of expanding economic opportunities; the operation of product and factor markets tends to marginalize their earnings or displace their skills. Once skills have been permanently marginalized or products displaced, the poor lack human and financial capital for adaptation.
Socioeconomic dualism	Sharply dualistic growth favors the rich and harms the relative and absolute position of the poor more than do lesser degrees of dualism (whether at very low or higher levels of development).	Concentration of growth in limited sectors or regions, in the presence of segmented labor markets characteristic of underdeveloped countries, aggravates displacement and marginalization of skills, thus contributing to pools of surplus labor.

---

[1] Irma Adelman and Cynthia Taft Morris. *Economic Growth and Social Equity in Developing Countries*. Stanford Univ. Press, Stanford, Calif. 1973.

[2] Irma Adelman and Cynthia Taft Morris. "Economic Growth and Impoverishment in the Middle of the Nineteenth Century." *Econ. Jour.* (forthcoming).

[3] Irma Adelman and Sherman Robinson. *Planning for Equitable Development*. Stanford Univ. Press, Stanford, Calif., (forthcoming).

Continued

Hypotheses concerning the impact of economic processes on income distribution—Continued

Variable or influence	Impact and supporting studies	Reasons for impact
Natural resources	A favorable ratio of agricultural land to population, with family-size holdings prevalent, restricts extreme poverty; abundant natural resources are associated with extreme concentration of income (A-M and Hist).	Availability of reserve of unappropriated agricultural land restricts numbers of extremely poor; presence of abundant natural resources fosters exploitation and appropriation of the benefits by small elites, foreign and indigenous.
Rate of population growth	Past patterns of population growth are an important determinant of resource-population ratios and thus a key initial condition determining extent of extreme poverty (Hist and A-M). In the medium run, population growth affects significantly the net balance of displacement and absorption processes (Hist and A-R). In the short run, population growth is insignificant to income distribution (A-R and A-M).	Extremely long lag (50 years or more) occurs before changes in rate of population growth affect extent of extreme poverty significantly; medium run effects depend on interaction of population and migration given the availability of resources.
Rural-urban migration	Rural-urban migration reduces rural poverty and increases urban poverty; up to a point, it reduces overall poverty (Hist and A-R).	Rural-urban migration in labor market reduces urban wages or increases urban unemployment, while decreasing rural underemployment or increasing wages of agricultural labor; in commodity markets, such migration shifts the terms of trade in favor of agriculture, thus favoring the more numerous and poorer rural low-income groups; extremely rapid migration leads unfavorable impact on the urban poor to dominate favorable rural effects.
Education	Widespread education and literacy are associated with larger share of income of middle quintile; no systematic association occurs with other features of income distribution (A-M and A-R).	Widespread education spreads ownership of human capital, distribution of wage income, and increases rural-urban migration; thereby shifting population to higher income areas and improving agricultural terms of trade.
Land tenure and holdings	Impact of change on the poor critically depends on distribution of landownership; prevalence of subsistence farming where land is abundant favors distribution of income; parcelization of land or marked concentration of ownership with cultivation by either landless laborers or subsistence tenants contributes directly to poverty; widespread owner-operation of commercial farms favors distribution of income, although it may not help the very poor (A-M and Hist).	Concentrated ownership where cultivators face lack of alternative employment permits high rate of appropriation of surplus product; opening up of new opportunities in presence of unequal abilities to respond widens inequality; processes conducive to dispossession of land or marginalization of economic activities tend to be irreversible.
Modernization of agricultural techniques	Increases in agricultural productivity tend to worsen the position of the rural poor, while benefiting better off farmers and the urban poor (A-R and Hist). No significant cross-sectional relationship exists.	Access to complementary resources (held only by better off farmers) is necessary for adoption of improvements; increases in output worsen terms of trade, harming rural poor and benefiting urban poor; no cross-sectional relationship exists due to interdependence between impacts of agricultural technology and land tenure.
Size of subsistence agriculture sector	Total absence of commercialization of agriculture favors distribution of income (Hist). Among countries with some commercialization, no systematic relationship exists between commercialization and distribution of income (Hist and A-M).	Impact on income distribution of commercialization of agriculture depends on land tenure and concentration of ownership. It also depends on the course of terms of trade; short-term and medium- or long-run impacts may diverge, depending on demand and supply elasticities.

Continued

Hypotheses concerning the impact of economic processes on income distribution—Continued

Variable or influence	Impact and supporting studies	Reasons for impact
Trade and industrialization strategy	More diversified labor-intensive exports are associated with a larger share of income to both the poor and the middle classes (A-M and A-R).	Agricultural terms of trade improve and rural-urban migration is more easily absorbed.
Level and change in industrialization	U-shaped relationship is evident in cross-sectional and historical studies. Major differences in level of industrialization are negatively associated with extent of poverty, but among more industrialized countries, poverty is greater in those at higher levels which industrialized most rapidly (A-M and Hist).	The relationship is complicated: interdependence exists between level and rate and threshold effect.
Market socialism	Share of income of top 5 percent is reduced; share of next 15 percent is increased (A-M and A-R). Overall growth rate and overall mean incomes are reduced but relative and absolute incomes of lowest 70 percent are higher in the medium run (A-R).	Profits and interest accruing to top 5 percent are reduced; technocrat and bureaucrat incomes are increased. Given nationalized firms which are less dynamic, absolute gains to the poor evident in the medium run are eventually eroded, but distributional gains increase.
Effectiveness of and improvements in financial institutions	No systematic association with income shares of the poor exists even though the effects are important for growth (A-M and A-R).	Spread of financial institutions even in rural areas tends to benefit primarily those who are better off. Improved access to credit for poor farmers is only beneficial when combined with improved access to technology and knowledge. Increases in investment, even in small-scale industry, tend to work through their impact on economic growth (relationship of latter to income distribution is complex).
Effectiveness of and improvements in tax systems	No systematic association with distribution in cross-sectional study exists (A-M). Little impact emerges in the model (A-R).	Structure of tax system influences who is poor and who is rich rather than how many; tax base is sufficiently low so that there is little scope for impact.

1. Intervention to improve the distribution of income is extremely difficult. The Korean policy model experiments emphasized both the stability of the size distribution of income despite policy interventions and also the ephemeral nature and inefficiency of most single-pronged antipoverty programs. The cross-sectional study indicated how few potentially effective policy instruments existed and what difficulties their very makeup posed for persons intent on purposive change.
2. Within a given set of structural conditions, it is strategy and process which determine the impacts of economic change on the poor. The historical study indicated, for example, that countries sharing a particular mix of poverty also shared a set of historical processes of change. The Korean policy model experiments showed that as long as policy interventions were tacked onto a given strategy which remained unchanged, the distribution of income tended to revert to the pattern it would have had in the absence of interventions.
3. Successful antipoverty policy does not merely entail choosing the right development strategy. In the cross-sectional study, for example, some policy variables proved important only when countries had achieved minimal levels of education and spread of economic modernization. The historical study also showed that the impact of economic change on the poor depends critically on the type of social structure and social responses to economic change—such as social constraints on population growth, the response of fertility and migration rates to changing economic opportunities, legal and customary barriers to the subdivision of land, arrangements for land tenure and holding, and the strength of extended family protection of the unemployed and underemployed.
4. Unbalanced growth strategies are bad for the poor. Historically, the concentration of growth in a few sectors or regions has had backwash effects that have accentuated the overall displacements of economic activities arising from commercialization and industrialization.
5. A systems or general equilibrium approach is required to design a strategy which predictably improves the position of the poor in the medium

> term. Both the historical and South Korean modeling studies emphasize the importance of indirect effects and dynamic interactions. Not infrequently, the indirect effects of an initial impact of change swamp the direct effects and even reverse their direction.

Taken together, results of the three studies underline the great difficulty which planners face in finding policy instruments that effectively bring about more equitable paths of economic growth.

This note summarizes the ERS Bicentennial Lecture delivered by Irma Adelman on May 17, 1976. Dr. Adelman is a professor in the Department of Economics at the University of Maryland.

REGIONAL POLICY: READINGS IN THEORY AND APPLICATIONS

John R. Friedmann and William Alonso, eds. The MIT Press, Cambridge, Mass. 808 + xxi pages. 1975.

The growing body of literature by regional scientists has spawned a number of anthologies around the world. Friedmann and Alonso's updating of their 1964 reader[1] is a useful addition. The editors culled 37 articles from 27 disparate sources and arranged them under 5 headings. Thirty-five authors are represented, including the editors. The readings are generally current; more than half were published during the 1970's.

The 15 readings comprising part 1, on concepts of space and development, fill nearly half the volume. These include basic references on micro and macro aspects of location, on motion within spatial systems, and on bases for growth. Several have the stature of classics, such as Losch's 1938 article, "The Nature of Economic Regions," and North's 1955 article, "Location Theory and Regional Economic Growth."

Seven articles in part 1 were in the first edition and two are in another reader.[2] All four articles on growth also appeared in the first edition; none of the vast growth literature since 1961 has been included in the current edition. A more up-to-date selection of readings on regional growth is available in *Regional Economics: Theory and Practice.*[3] Theoretical aspects of spatial concepts are probed more deeply in the reader by Dean, *et al*, and a broader collection of readings on location is available in the reader by Smith, Taaffe, and King.[4] The duplication of this first half, not only with the first edition but also with other available readers, suggests that the success of this new volume must depend on what the editors accomplish in the second half.

The six readings in part 2, on the role of cities, are introduced with the observation that:

> Cities play a preeminent role as the nodal points in [the structure of space], as the holders of power, and as the seedbeds of change. They are also seats of learning, engines of production, and confluences of power and control. In developed countries . . . the cities relate primarily each to each . . . .

Several of the subsequent articles consider whether cities are too large or too alienating, whether developing countries are "overurbanized," and whether urban problems are likely to develop into political upheaval. However, the several authors manage, by and large, to find satisfactory answers to these questions. Generally, they appear to support the editors' urban point of view. An exception is the 1954 article by anthropologists Redfield and Singer on the cultural role of cities. They challenge the view that there is a one-way flow from urban to rural areas, and they trace vital interactions and feedbacks between the two sectors. This urban bias did not exist in the first edition, which contained four articles on "Problems of the Rural Periphery."

The eight readings in part 3, on formulating issues in regional policy, are more what one would expect from the book's title. It is here that this volume of readings comes into its own. The authors grapple with means for achieving a number of ends, including the geographic distribution of population and income. Conflicts are probed such as the trade-off between equity and efficiency, and between improved welfare and the status quo. Methodological approaches are examined. The scholarly article by Darwent clears up much confusion about the terms "growth pole" and "growth center." The author relates these concepts not only to economic growth theory, but also to applied regional policy.

These policy readings reflect the fact that the regional dimensions of national purpose in the United States are not well articulated, and that strong regional impacts of national policy tend to be experienced as unintended side effects rather than as purposive ends.

The seven readings in part 4 deliver the "applications" promised in the subtitle of the book. The authors present case studies of regional planning in Turkey, Israel, Italy, Spain, and France. In view of the current debate on centralization versus decentralization in regional planning, Tekeli's dramatic description of central control of regional allocation in the Ottoman Empire is fascinating. Tekeli takes us back to the 16th century when the Empire was at its height. Central policy spelled out the crop to be grown in each region and the allocation of the product among cities; a central documentation system specified laborers by region, skill, and name. This degree of meticulous and broad detail occurred long before the advent of computers and linear programming.

The idea presented by one of the editors, that regional science is an increasingly international field, is borne out not only by the international flavor of part 4, but is also evidenced throughout the volume. More than half the articles apply specifically to countries other than the United States, or depend heavily on international experience. One out of every five authors listed in the introductory material has an address outside the United States.

---

[1] Friedmann, John R. and William Alonso, eds. *Regional Development and Planning: A Reader.* Cambridge, The MIT Press, 1964.

[2] Needleman, Lionel, ed. *Regional Analysis: Selected Readings.* Baltimore, Penguin Books, Inc., 1968.

[3] McKee, David L., Robert D. Dean, and William H. Leahy, eds. *Regional Economics: Theory and Practice.* New York, The Free Press, 1970.

[4] Dean, Robert D., William H. Leahy, and David L. McKee, eds. *Spatial Economic Theory.* New York, The Free Press, 1970.
Smith, Robert H. T., Edward J. Taaffe, and Leslie King. *Readings in Economic Geography: The Location of Economic Activity.* Chicago, Rand McNally, 1968.

Part 5 is a bibliographical essay by editor Friedmann. He reviews progress in theoretical knowledge, policy analysis, empirical experience, and regional planning, as reported in the literature during the decade following the first edition.

Werner Hirsch, in the December 1965 *American Economic Review*, labeled the first edition of Friedmann and Alonso's reader a potpourri. Hirsch observed the first volume had little structure or discipline. The editors have not changed their approach much in 10 years. Introductions to each section do help users find their way through the loose structure and overlapping categories. An index would have helped even more.

This new potpourri contains many excellent ingredients and a few of doubtful value. It omits many fine articles that would have improved the mixture. But, on balance, the readings form a worthwhile collection for regional scientists.

Clark Edwards, Economist
Economic Development Division

A NOTE ON THE RECENT DECLINE IN THE AVERAGE AGE OF FARMERS

As soon as the number of farms began to decline in the United States—which was at the time of World War I—the average age of farmers began to rise. This change was very logical, for an occupation that decreases in number of workers typically does so mainly because young persons fail to enter that line of work. Thus, the average age of those remaining gradually increases. The average age of all farm operators rose from 43.5 years in 1910 to 51.3 years in 1965. In the same period, the proportion of farmers who were less than 35 years old dropped from 29 percent to 11.5 percent.

Adjustments in agriculture to the declining need for workers brought about by mechanization and other labor-saving developments have stretched out over two generations since World War I. In public debate and commentary on American agriculture in the 1950's and 1960's, it was common to encounter concern over the advancing age of farmers. Where would it all end? Who would take over when today's farmers died or retired? But the long decline in number of young farmers seems to have ended since 1970.

Yearly statistics collected by the Bureau of the Census from national sample surveys show that the median age of persons self-employed in agriculture reached a peak of 53.1 years in 1970. This figure differs somewhat from the data cited above in that it includes a small percentage of people who are not farm operators (such as those in veterinary work, crop dusting, and cotton ginning), and it includes only persons whose sole or principal job is in agriculture. But these differences do not affect the validity of the conclusion.

From 1970 to 1975, the average age of self-employed persons in agriculture dropped from 53.1 to 50.4 years. The number of self-employed agricultural workers under 35 years old rose from 265,000 to 358,000, a gain of 35 percent. At the same time, the number aged 60 years or over dropped from 601,000 to 461,000, a decrease of 23 percent.

Persons self-employed in agriculture by age, annual averages, 1967-75[a]

Age categories	1967	1968	1969	1970	1971	1972	1973	1974	1975
				Number (thousands)[b]					
Total	1,996	1,985	1,896	1,810	1,748	1,789	1,776	1,752	1,715
16-34 years	284	295	272	265	274	295	298	329	358
35-59 years	1,124	1,097	1,011	944	926	932	929	914	896
60 years and over	590	593	613	601	549	562	548	510	461
Median age	52.2	52.2	52.9	53.1	52.6	52.3	52.3	51.5	50.4
				Percentage distribution					
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
16-34 years	14.2	14.9	14.4	14.6	15.7	16.5	16.8	18.8	20.9
35-59 years	56.3	55.3	53.3	52.2	53.0	52.1	52.3	52.2	52.2
60 years and over	29.6	29.9	32.3	33.2	31.4	31.4	30.9	29.1	26.9

[a] Persons 16 years and over. [b] Numbers are rounded to nearest thousands without adjustment to group totals.

Source: *Employment and Earnings*, Bur. Labor Statistics, U.S. Dept. Labor. 1968-76, January issues.

The figures are mute, of course. One can only conjecture about the circumstances that have finally halted the rise in the age of farmers at this time. But it was known that the aging process could not go on forever. Agriculture was not disappearing. Rather, it was simply going through a transition, and the day had to come when the number of workers reached some stabilization and the age composition began to normalize.

Talks with local agricultural officials have indicated for several years that the new trend is recognized in prime farming areas of the Midwest. In this writer's experience, the most commonly suggested reasons offered are the improved profit picture in farming, a change in attitude of farm youth toward agriculture and rural life compared with city life, and the entry of urban-reared young adults into farming, based partly on philosophical considerations. We lack systematic evidence and specific details on the trend. It can also be said that the trend is currently helped by the large number of young adults available—the "baby boom" children.

Whatever the details to be filled in, the existence of the new trend is definite. If there were two demographic articles of faith in modern times for those engaged in rural policy or research, they were the advancing average age of farmers and the high rate of rural-to-urban migration. Now, both of these trends have changed in the 1970's.

Calvin Beale
Leader, Population Studies
Economic Development Division

TECHNOLOGY ASSESSMENT WORKSHOP

Technological advance has been considered a source of progress over the centuries. In recent times, we have become increasingly aware of the indirect and delayed impacts of such advance on society. Frequently, they can be undesired and unintended. A systematic study of these impacts is coming to be known as "technological assessment."

ERS sponsored a workshop on technology assessment (TA) at the National 4-H Center in Chevy Chase, Md., April 20-22, 1976. There were three main purposes:

- To brief researchers on new concepts and developments in technology assessment;
- To review the status of ERS research and methods in relation to new technological concepts;
- To determine the needs for expansion of ERS work in technology assessment.

Although most of the participants came from ERS, representatives also participated from the Agricultural Research Service (ARS), Foreign Agriculture Service, Congressional Research Service (CRS), Office of Technology Assessment (OTA), Midwest Research Institute, and National Science Foundation (NSF).

In the opening paper, Quentin M. West (ERS) gave two major reasons why ERS should expand its work in technology assessment: (1) to perform the ERS mission better and (2) to enhance contributions of new technology to societal objectives of increasing productivity while reducing adverse environmental, physical, and social impacts. Joseph F. Coates (OTA) followed this presentation with a discussion of the inevitability of TA, and explained how technology can become part of the problem instead of the solution. Anticipating future impacts of current technological advance, he said, is as important for what he called "social technology" as it is for physical technology. Papers by Marshall E. Miller and W. B. Back (ERS) and Glen E. VandenBerg (ARS) focused on the interdependence of research and technology assessment and on how TA can contribute to research planning. TA's reliability depends largely on the soundness of a knowledge base developed through research. In turn, technology assessment contributes information useful for making decisions about research resource allocation.

In his account of major technological developments and their impacts over the past two centuries, Wayne D. Rasmussen (ERS) reminded us of agriculture's neglect of the social effects of technology. He constructed a hypothetical chain of events that showed how Eli Whitney's invention of the cotton gin could have contributed, indirectly, to the Civil War nearly 70 years later. The cotton gin was a prerequisite for the profitable expansion of cotton production. Without such expansion, slavery might have disappeared. Without slavery, there might not have been a Civil War.

Richard G. Stuby (ERS) made two main points in his paper on conceptualizing and measuring social impacts of technology. First, social impacts are not what remains when the economic factors are removed; rather, they encompass economic impacts. Second, the usual distinctions between qualitative and quantitative measurement obscure the differing, significant mathematical properties of phenomena.

On the second day of the workshop, speakers addressed the status of agricultural and rural technology assessment. Topics included the flue-cured tobacco harvester, the new 4-wheel drive tractor, integrated hog farming, twinning in beef cattle, coal mining in the Great Plains, boll weevil management, and irrigation technology. Midwest Research Institute, under contract from NSF, conducted the study on integrated hog farming. A major finding was that hog production will probably not be integrated as rapidly as poultry production was. The other studies were carried out by ERS researchers.

These papers showed where we stand in TA—the progress made and in the making. But, as a whole, they exhibited a major weakness of agricultural TA to date. It continues to stress feasibility of new technology for adoption within the food and fiber industries instead of addressing the broad social impacts of such technology.

Participants examined the utility of various TA methods on the third day of the workshop. Jean M. Johnson (NSF) presented a paper on "Role of Models in Technology Assessment." Papers followed on simulation, scenario writing, the Delphi technique, statistical methods, input-output analysis, parametric and linear programming, forecasting adoption, and benefit-cost analysis. Most papers included examples or illustrations of methods in studies completed or in progress. In the paper on Delphi, Douglas McNiel and Yao-Chi Lu (ERS) described how the technique was being used to forecast new technologies in meatpacking. Barry Carr (CRS), in his paper on scenario writing, gave many examples of the use of this approach in TA, including studies on alternative futures of U.S. agriculture and on minimum tillage. Apparently, no unique methods exist, judging from the papers, of making technology assessments. Rather, the existing research methods have varying degrees of utility within the several operations of technology assessments.

Proceedings of the workshop will be published and will include reviews and appraisals by two groups.

W. B. Back
Agricultural Economist
National Economic Analysis Division

U.S. GOVERNMENT PRINTING OFFICE : 1976 O - 222-188

Suggestions for Submitting Manuscripts for *Agricultural Economics Research*

Contributors can expedite reviewing and printing of their papers by doing these things:

1. **SOURCE.** Indicate in a memorandum how the material submitted is related to the economic research program of the U.S. Department of Agriculture and its cooperating agencies. State your own connection with the program.
2. **CLEARANCE.** Obtain any approval required in your own agency or institution before sending your manuscript to one of the editors of *Agricultural Economics Research*. Attach a copy of such approval to the manuscript.
3. **ABSTRACT.** Include an abstract and at least three keywords when you submit your article. The abstract should not exceed 100 words.
4. **NUMBER OF COPIES.** Submit three good copies.
5. **TYPING.** Double space everything, including abstract and footnotes.
6. **FOOTNOTES.** Number consecutively throughout the paper.
7. **REFERENCES.** Check all references carefully for accuracy and completeness.
8. **CHARTS AND OTHER ARTWORK.** Use charts sparingly for best effect. Keep design as simple as possible to improve communication. Submit all artwork in draft rather than final form, accompanied by neatly prepared pages with essential data for replotting. Complex or detailed charts and other artwork are usually best suited to full page or 2 page treatment (final image size will be 6-5/8" x 4-3/8"). Simpler charts should be designed to fit half page (final image size of 6-5/8" x 4-3/8") or quarter page (final image size of 3-3/8" x 4-3/8").
9. **FINAL TYPING.** Two good copies, double spaced, will be required of final copies of edited and revised manuscripts accepted for publication.

UNITED STATES DEPARTMENT OF AGRICULTURE
WASHINGTON, D.C. 20250

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF
AGRICULTURE
FIRST CLASS
AGR 101

U.S.MAIL

AGER HOLE 453E ISSDUE002R
ERLING HOLE
6453 OAKWOOD DR
FALLS CHURCH VA 22041

AGRICULTURAL ECONOMICS
RESEARCH

The Secretary of Agriculture has determined that the publication of this periodical is necessary in the transaction of the public business required by law of this Department. Use of funds for printing this periodical has been approved by the Director, Office of Management and Budget, through February 28, 1980.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. $1 a single copy, $3.85 a year domestic, $4.85 foreign.